

PENGOLAHAN DAN ANALISIS DATA DENGAN SPSS

by Sudung Nainggolan

Submission date: 11-Feb-2019 01:58PM (UTC+0700)

Submission ID: 1076255033

File name: Bahan_Buku_SPSS_Edit_0601.docx (6.56M)

Word count: 30942

Character count: 183530

PENGOLAHAN DAN ANALISIS DATA DENGAN SPSS

33

Dr. Sudung Nainggolan, BSc., MHSc.

Departemen Ilmu Kesehatan Masyarakat/Kedokteran Komunitas

Fakultas Kedokteran

Universitas Kristen Indonesia

ISBN:

Cetakan pertama

Penerbit Nusa Hak Cipta

Penerbit

Editor: Budi Purwoto

KATA PENGANTAR

Biostatistik merupakan salah satu cabang dari Ilmu Statistik dalam bidang kesehatan dan biologi. Karena perannya begitu luas, biostatistik, bukan saja menjadi bagian dari kurikulum di Fakultas Kedokteran, Ilmu Kesehatan Masyarakat, Keperawatan dan lain sebagainya tetapi juga diajarkan pada kursus-kursus akademik dan profesional. Di masa lalu para mahasiswa dan peneliti menghabiskan waktu yang cukup lama dalam menyelesaikan tugas statistik secara manual. Namun kini dengan munculnya komputer, beberapa program dibuat tersedia untuk menganalisis data statistik. Salah satunya adalah SPSS (Statistical Package for Social Science).

Pertama kali penulis terpapar dengan SPSS(X) pada tahun 1984 ketika sebagai mahasiswa di Departemen Biostatistik - School of Hygiene and Public Health Johns Hopkins University, Baltimore - USA. Tidak seperti saat ini paket SPSS sudah tersedia untuk dijalankan di Personal Computer atau Laptop/Notebook, penulis sebelumnya mengerjakan tugas-tugas akademik dengan bahasa *Programming SPSS(X)*[®] yang cukup kompleks pada platform komputer mainframe.

Berdasarkan pengalaman dalam sepuluh tahun terakhir penulis mengajarkan SPSS bagi pemula, dalam membimbing khususnya para mahasiswa dalam penelitian berikut analisis statistik mendorong minat dan keyakinan untuk menulis buku panduan belajar SPSS secara mandiri. Kami masih menggunakan SPSS Versi 17 (Licensed) di Universitas ditempat kami mengajar.

Adapun ruang lingkup buku ini adalah untuk memperkenalkan pembaca kepada SPSS Windows dan untuk memungkinkan mereka memasukkan dan memformat data, menjalankan analisis, membuat berbagai jenis diagram dan grafik dan menafsirkan data. Buku ini disiapkan untuk digunakan dalam pengajaran biostatistik di perguruan tinggi dan bagi mereka yang bekerja secara mandiri dalam penelitian untuk analisis dan interpretasi data.

Buku ini ditulis dengan sederhana. Subyek diatur dalam bab dan bagian, dinomori oleh sistem penomoran. Semua bab ditulis seperti tutorial. Setiap bab memiliki instruksi yang memandu mahasiswa melalui serangkaian latihan, dan grafik yang menunjukkan bagaimana tampilan layar di berbagai langkah dalam proses.

Buku ini memiliki delapan bab. Secara ringkas, Bab diawali dengan pengenalan jargon-jargon SPSS diikuti penjelasan Statistik Deskriptif. Bab-bab selanjutnya berisi materi khusus bab dengan latihan. Bab 3 secara eksklusif membahas cara memproduksi grafik yang serbaguna seperti diagram batang berkelompok, dan lain-lain dengan bantuan *Chart Builder* dan grafik Interaktif. Bab tentang membandingkan rata-rata, analisis varians, korelasi, regresi dan chi-square ditulis dengan cara yang sangat sederhana dengan contoh-contoh spesifik, untuk memungkinkan pembaca dapat memahami konsep dan melakukan analisis dengan mudah, dan menafsirkan hasilnya.

Latihan soal-soal pada setiap bab dimasukkan khusus untuk memungkinkan pemula untuk memiliki keahlian dalam mengerjakan secara pribadi. Sehingga secara tidak langsung pembaca digiring untuk memahami latar belakang teori yang mendasari dalam contoh-contoh kasus yang tersedia. Untuk itu, kami juga menyertakan glosarium untuk bantuan rujukan.

Ucapan terima kasih kepada para penulis berbagai buku tentang SPSS yang telah kami rujuk untuk menulis buku ini. Secara khusus kepada Profesor Rajathi dan Profesor Chandran yang telah menulis banyak tentang contoh-contoh kasus di bidang kedokteran dan biologi.

Kami berharap buku ini akan dapat membantu para pembaca dalam melakukan analisis dengan SPSS. Jika Anda ingin memberi saran, memperbaiki kesalahan, atau memberi kami umpan balik, dengan senang hati kami persilahkan. Kirimkan saran dan kritik Anda ke sudung.nainggolan@fkuki.ac.id, untuk memungkinkan kami meningkatkan isi dalam edisi berikutnya.

Sudung Nainggolan

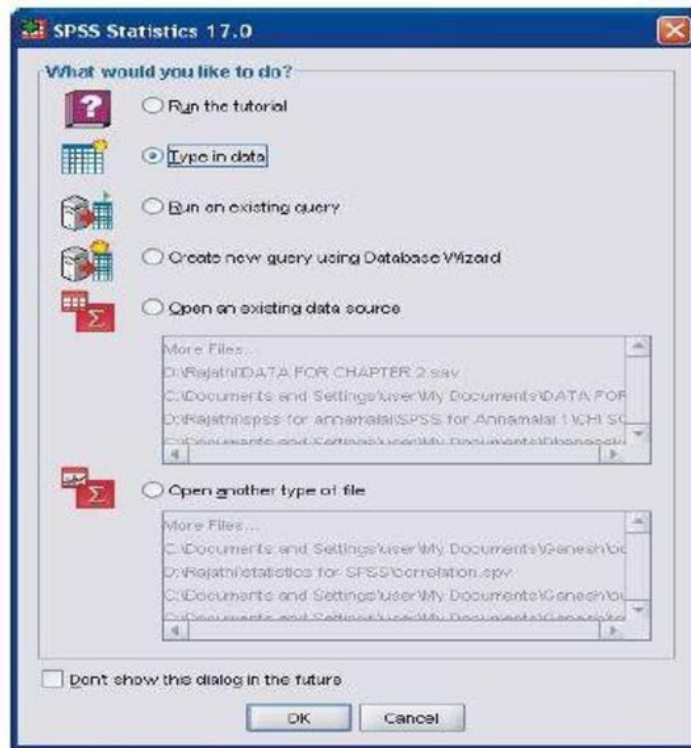
Bab 1

FILE DATA SPSS

MEMBUKA FILE DATA DI SPSS

Ada beberapa cara membuka file data di SPSS. Salah satu caranya adalah dengan mengklik ikon SPSS. Jendela pengantar akan muncul dengan judul statistik SPSS 17.0 (Gambar 1.1). Baca dan klik tombol berlabel **Type in data** lalu klik OK.

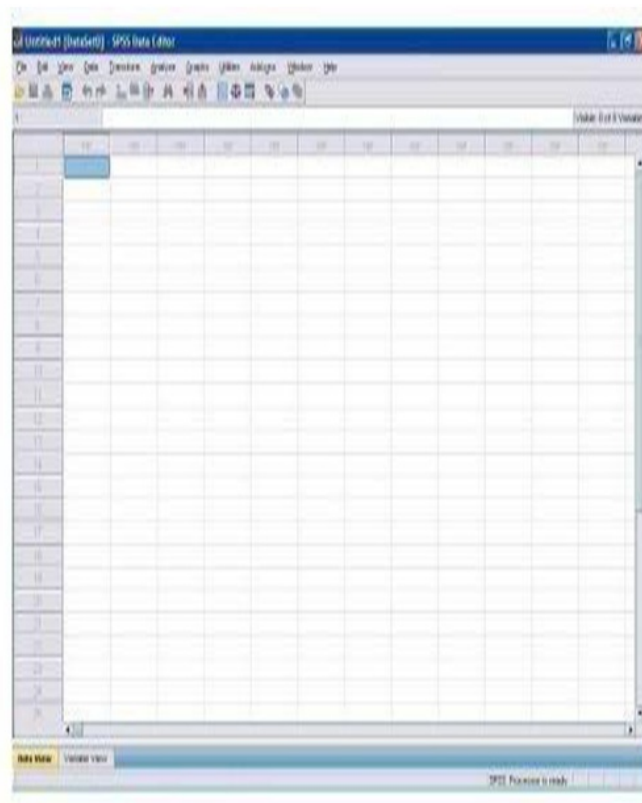
Sekarang **Data Editor** muncul **Variable View** di bawah layar Di kaki Data Editor, **Data View** muncul bersama dengan Variabel View. Ada juga **status bar** yang menunjukkan garis, SPSS Processor sudah siap. Kita harus memeriksa ini saat bekerja. Data diketik langsung di file data SPSS yang sudah dibuat di Data Editor. Data juga dapat diimpor dari Excel dan STATISTICA. Dalam kumpulan data SPSS, setiap baris hanya mewakili satu kasus dan setiap kolom mewakili variabel atau karakter kasus yang diukur. Sebelum memasukkan data di Data Editor, penting untuk memahami istilah yang digunakan dalam data editor.



Gambar 1.1 Membuka file data di SPSS

DATA EDITOR

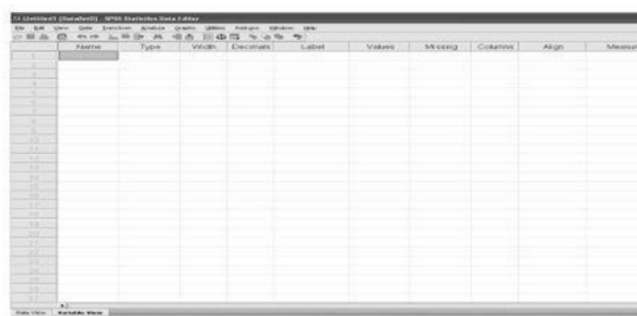
Data Editor SPSS memiliki dua spreadsheet yang tersusun. Salah satunya adalah Data View, di mana data baru dimasukkan dan yang lainnya adalah Variabel View yang berisi nama dan detail variabel data (Gambar 1.2).



Gambar 1.2 Data Editor SPSS

VARIABEL VIEW

Untuk mendapatkan Variabel View, klik Variabel View di bagian bawah ki⁴⁸ jendela. Sekarang lembar data muncul dengan judul Variable View seperti pada Gambar 1.3.



Gambar 1.3 Variabel View

Lembar data Variabel View memiliki 10 kolom yaitu:

1. **Name.** Ini adalah karakter string (biasanya, huruf dan spasi, dan kadang-kadang digit). Tampaknya di atas kolom di Tampilan Data tetapi tidak di output. Ini adalah tampilan singkat yang hanya muncul dalam tampilan data. Ini harus menjadi urutan terus menerus tanpa spasi. Meskipun 64 huruf dapat dimasukkan, diinginkan untuk membuatnya tetap pendek. Ini bisa menjadi campuran kasus.

2. **Type.** Ini menerima delapan jenis variabel yang berbeda. Dua yang penting adalah numerik, yaitu: angka dengan titik desimal dan string, atau apa saja terkait karakter non-numerik.

3. **Width.** Ini adalah Width dari variabel. Pengaturan default untuk Width variabel adalah 8. Tapi bisa diubah dengan memilih Edit, mengklik opsi dan kemudian memilih Data.

4. **Decimal.** Ini adalah jumlah desimal yang akan ditampilkan dalam Tampilan Data. Pengaturan standar menampilkan 2 desimal. Jika diperlukan bisa diubah dengan mengklik dua kali pada panah ke atas atau bawah.

5. **Label.** Label adalah frasa yang bermakna dengan spasi di antara kata-kata. Ini menggambarkan variabel dan juga muncul di output. Penting untuk menetapkan label yang memberi arti untuk setiap variabel.

6. **Values.** Kolom ini dimaksudkan untuk pengelompokan variabel yang memberikan kunci untuk arti angka kode. Kotak dialog values dibuka dengan mengklik area abu-abu. Values dan labelnya diberikan dalam kotak dialog values.

7. **Missing.** Ini menentukan nilai/angka/data yang missing dalam satu set data.

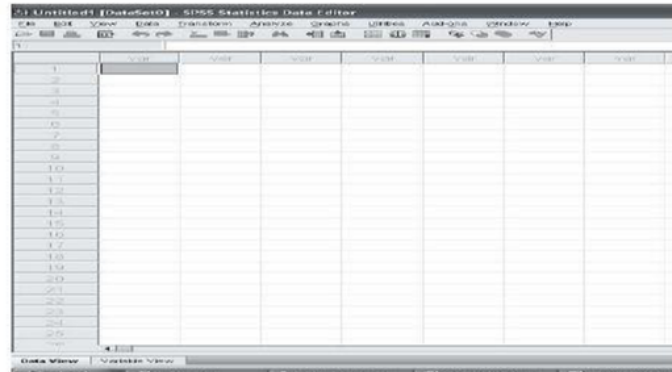
8. **Columns.** Menunjukkan Width semua variabel yang muncul di Data View.

9. **Align** Kolom ini menentukan apakah data yang tersisa, kanan atau tengah sejajar. Standarnya adalah rata kanan.

10. **Measures.** Ini menjelaskan jenis skala pengukuran, apakah data berada di skala, ordinal atau nominal.

DATA VIEW

Data View dari Data Editor saat diakses memberikan nama variabel di bagian atas kolom yang berlabel di Variabel View. Sisa dari kolom berisi nama default variable, menunjukkan bahwa kolom-kolom ini belum diberi label dengan variabel tertentu (Gambar 1.4).



Gambar 1.4 Data View

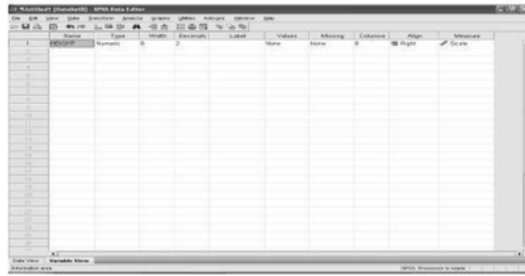
MEMASUKKAN DATA MENJADI DATA EDITOR

Penamaan Variabel dalam Variabel View.

Untuk penamaan variabel dan propertinya, klik Variabel View. Penamaan variabel diserahkan kepada preferensi individu. Tetapi Anda harus ingat untuk menggunakan nama variabel yang sama persis untuk data set.

Penamaan Variabel Numerik dalam Variabel View

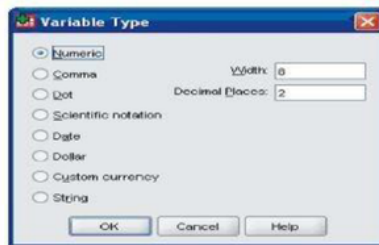
Langkah 1 Masukkan nama variabel di bawah **Name** dalam Variabel View. Misalnya, jika Anda ingin memasukkan tinggi individu di kelas, ketik HEIGHT, di bawah Name (Gambar 1.5). (Harus diingat bahwa itu adalah berurutan tanpa spasi di antara karakter. Apa pun yang kita ketikkan, itu muncul di atas kolom di Data View tetapi tidak di output).



Gambar 1.5 Penamaan variabel numerik dalam Variabel Lihat

Langkah 2 Selanjutnya ke Type, dan klik kanan di mana saja di sel kolom Type. Kotak dialog Variable Type terbuka (Gambar 1.6).

Langkah 3 Pertahankan format Numeric (tipe default). Tetapkan dan masukkan Width dan **Decimal Places**. Karena variabelnya adalah "Height" maka kita hanya memilih dua desimal.

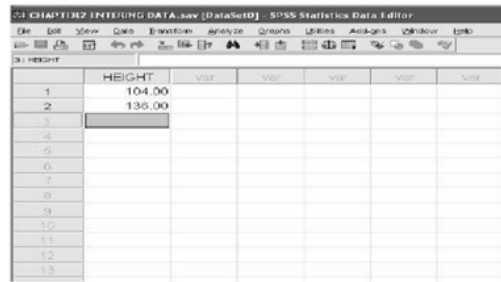


Gambar 1.6 Memilih tipe variabel dan decimal places

Langkah 4 Di bawah kolom Label, jelaskan variabel dengan sebutan atau kalimat. Ketiklah dengan benar karena ini muncul di output. Misalnya, kita dapat mengetik sebagai "Tinggi Mahasiswa Semester 8".

Langkah 5 Karena Height adalah variabel numerik kita perlu memberikan unit ekspresi.

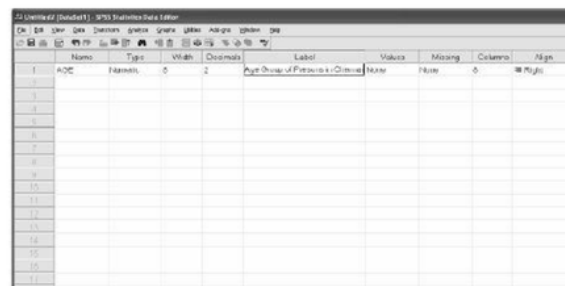
Langkah 6 Selanjutnya kita dapat membuka Data View dengan mengklik yang sama di bagian bawah. Data View muncul seperti yang diberikan pada Gambar 1.7 dengan label variabel sebagai "HEIGHT". Sekarang kita sudah bisa mulai mengetik data.



Gambar 1.7 Memasukkan data dalam Data View

Memasukkan data untuk variabel dikelompokkan atau kategori dan penamaan variabel dikelompokkan atau kategori dalam Variabel View. Jika kita tertarik untuk mengetahui, misalnya, pentingnya perbedaan tekanan darah di antara kelompok usia dalam suatu populasi, kita dapat memasukkan variabel seperti yang dijelaskan dalam paragraf berikut.

Ukur tekanan sistolik dalam mmHg untuk kelompok usia yang berbeda dan kategorikan usia (variabel) menjadi muda, dewasa dan tua, sebelum memasukkan data. Sekarang, klik pada tab Variabel View di kaki Data Editor. Masukkan nama variabel sebagai "usia", lanjut ke kolom Type, tetap dengan Numeric Format (default) dan tentukan Width dan Decimal Places. Jelaskan variabel di bawah kolom Label (seperti yang Anda inginkan muncul di output. Misalnya, usia dalam tahun di Kota X). Teruskan ke kolom Value, klik pada area abu-abu, jendela pop up akan terbuka (Gambar 1.8).



Gambar 1.8 Menamai variabel kategori atau pengelompokan dalam variabel view

Anda dapat memberikan nomor kode dalam Value dan kunci untuk kode di Label. Misalnya, ketik "1" dalam Kotak Value dan "young age" (usia 1-18), di kotak Label" sama dengan label "2" "adult" (usia 19-50) dan "3" sebagai "old" (usia 51 tahun dan di atas) (Gambar 1.9).



Gambar 1.9 Menamai variabel kategori atau pengelompokan dalam kotak dialog label

Kemudian ketik variabel kedua sebagai "Blood Pressure" di bawah Name dan selesaikan sisanya. Kemudian klik OK untuk kembali ke Data View. Ketik data di bawah kepala tertentu (Gambar 2.10).

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	AGE	Numeric	8	2	Age Group of Respondents	None	None	8	Right
2	BLOODPR	Numeric	8	2	Blood pressure in mmHg	None	None	8	Right

Gambar 1.10 Variabel yang disebutkan dalam Variabel View dengan perincian untuk dua variabel

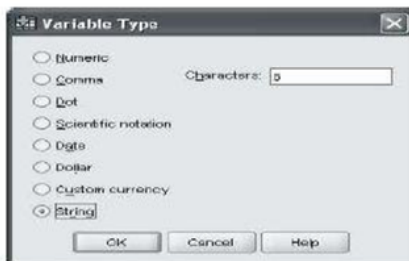
Penamaan variabel kualitatif dalam Variable View.

Jika Anda ingin mengetik golongan darah mahasiswa di kelas, klik pada Variabel View, ketik Bloodgpn pada Name (Gambar 1.11).

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	GENDER	String	8	0	Gender	None	None	8	Left
2	BLOODGPN	String	8	0	Blood Group	None	None	8	Left

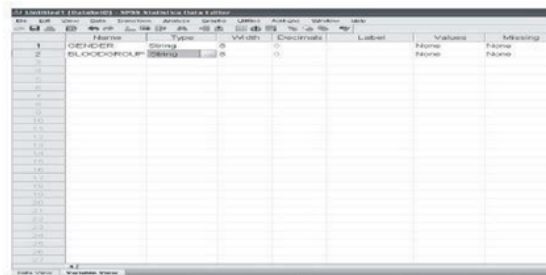
Gambar 1.11 Penamaan Variabel Kualitatif dalam Variabel View

Atribut atau variabel kualitatif, dinamai dalam Variabel View. Lanjut ke Type, klik kanan di mana saja di sel di bawah kolom Type, kotak dialog Variabel muncul. Pilih tombol String dan kemudian klik OK untuk kembali ke variabel view. Label variabel (seperti pada contoh sebelumnya). Tidak perlu menamai kolom Value karena Anda telah memilih variabel String (Gambar 1.12).



Gambar 1.12 Memilih String dari kotak dialog Variable Type (Untuk variabel kualitatif dalam variabel view)

Ketik variabel-variabel kualitatif apapun seperti pada Gambar 1.13.



Gambar 1.13 Dua variabel kualitatif yang disebut dalam Variabel View

Setelah menentukan semua variabel dan karakteristiknya, klik tab Data View di kaki Variabel View untuk membuka file data. Masukkan data Anda berdasarkan kasus.

Memasukkan data dalam Data View

Setelah spesifikasi dimasukkan ke Variable View, klik tab Data View di bagian bawah Variable View. Ketika Tampilan Data terbuka, nama variabel akan terlihat di judul kolom (Gambar 1.14).

	GENDER	BLOODGROUP				
1	Male	A				
2	Male	B				
3	Male	A				
4	Male	B				
5	Male	C				
6	Male	AB				
7	Male	C				
8	Male	AB				
9	Male	B				
10	Male	A				
11	Male	A				
12	Female	B				
13	Female	C				
14	Female	AB				
15	Female	C				
16	Female	AB				
17	Female	B				

Gambar 1.14 Data dimasukkan untuk dua variabel Kualitatif dalam Data View

(Catatan: Periksa di bagian bawah jendela tampilan data, di mana bar horizontal muncul dengan pesan: Prosesor SPSS sudah siap. Tanda horizontal ini dikenal sebagai "Status Bar". Nama ini disebut demikian untuk memberitahukan bahwa prosedur SPSS telah siap untuk dimulai).

Isi sel yang Anda ketik juga ditampilkan di area putih yang dikenal sebagai **Cell Editor** tepat di atas judul kolom. Nilai-nilai dalam sel dapat diubah dengan mengklik di editor sel, lalu dengan memilih Value sekarang dan menggantinya dengan nilai baru.

Nilai baru akan muncul di tempatnya. Adalah mungkin untuk menyorot sel atau seluruh blok sel atau seluruh baris atau kolom. Ini akan membantu Anda menyalin nilai dari satu kolom dan menempelnya ke kolom lainnya. Demikian pula entri dapat dihapus dengan menekan tombol hapus.

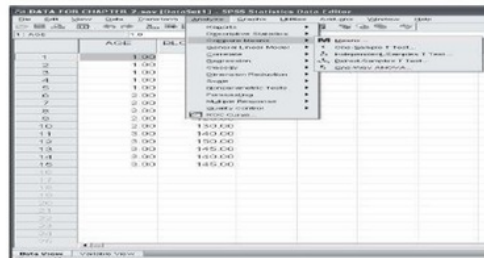
MENYIMPAN FILE DATA

Setelah entri data selesai dan diperiksa keakuratannya, pilih File dari menu utama, dan klik **Save as**. Kotak dialog Save as akan terbuka. Tentukan lokasi penyimpanan yang sesuai untuk file ini (seperti disk C atau D). Selalu ada baiknya menyimpan data dalam folder yang dibuat sebelumnya dengan nama seperti latihan SPSS 17.0. Ketik nama file di kotak **File Name** dan klik Save. Tutup SPSS dan jendela terbuka lainnya sebelum keluar dari komputer. Jika Anda tidak melakukan ini, Anda harus memberikan nama file ketika Anda mengakhiri atau menutup SPSS.

ANALISIS STATISTIK

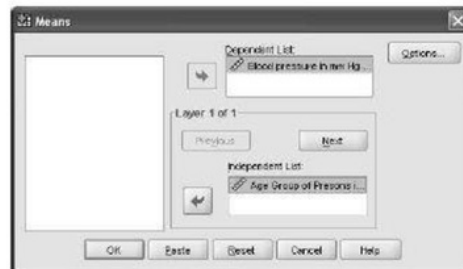
Kita mulai dengan contoh menghitung untuk variabel kategori yang terdapat pada Gambar 1.10.

Langkah 1 Dari menu utama, pilih **Analyze**. Dari menu **drop-down**, pilih **Compute Means** dan kemudian Rata-rata hitung (Means) (Gambar 1.15).



Gambar 1.15 Memilih Mean dari menu utama di Data Editor

Langkah 2 Membuka kotak dialog Means. Di panel kiri nama variabel terlihat (Gambar 1.16).



Gambar 1.16 Memilih dan mentransfer daftar variabel ke dependen dan independen di kotak Dialog Means

Langkah 3 Klik pada variabel, "Age group of person", lalu pada panah menunjuk ke **Independent List** untuk mentransfer variabel ke kotak yang sesuai. Demikian pula, klik pada variabel, "Blood Pressure" kemudian pada panah menunjuk ke kotak **Dependent list text box** untuk mentransfer variabel.

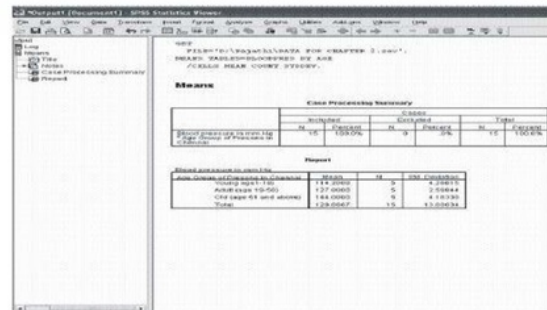
Langkah 4 Klik Opsi untuk membuka **Means:Options** Kotak dialog (Gambar 1.17).



Gambar 1.17 Memilih dan mentransfer deskripsi (seperti rata-rata hitung ...) di bawah Statistik ke Sel Statistik

Langkah 5 Transfer Rata-rata hitung, Jumlah Kasus dan Standar Deviasi dari panel kiri (Statistik) ke panel kanan (Sel Statistik).

Langkah 6 Klik Lanjutkan dan kemudian OK untuk menjalankan analisis. Hasilnya akan muncul di jendela baru yang disebut Output Viewer (Gambar 1.18).



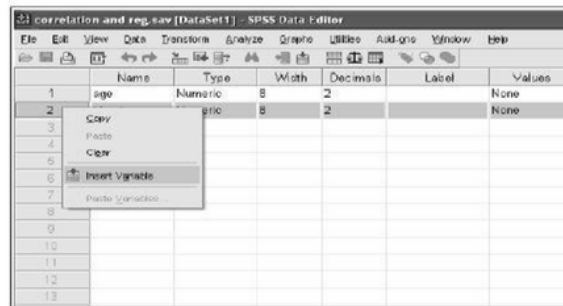
Gambar 1.18 Tampilan Output SPSS menunjukkan daftar item Output di panel kiri dan tabel output di panel kanan

MENGEDIT DAN "MANIPULASI" DATA

Setelah memasukkan data, selanjutnya bisa diedit atau dimanipulasi seperti memasukkan variabel baru, mengatur ulang urutan variabel dalam Variabel View atau mengubah jenis variabel.

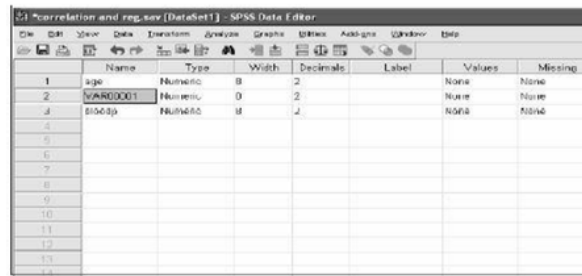
MENYISIPKAN VARIABEL BARU

Variabel tambahan dapat disisipkan dalam Variabel View dengan menyorot setiap baris dengan mengklik sel abu-abu di sebelah kiri dan memilih **Insert variable** (Gambar 1.19).



Gambar 1.19 Memasukkan variabel baru dalam Variabel View

Variabel baru, dengan nama default "VAR00001" akan muncul di atas baris yang telah disorot (Gambar 1.20).



Gambar 1.20 Variabel baru disisipkan di antara yang lama

Sekarang ketik nama variabel dalam Variabel View.

MENGATUR ULANG VARIABEL DALAM VARIABEL VIEW

Seperti memasukkan variabel baru, urutan variabel dalam Data View dapat diubah. Untuk itu klik Variabel View lalu klik kotak abu-abu di sebelah kiri variabel skor untuk menyorot seluruh baris. Dengan menahan tombol kiri mouse ke bawah, seret (drag) penunjuk layar ke atas. Garis merah akan muncul di atas baris grup. Saat melepaskan tombol mouse, skor variabel akan segera muncul di bawah huruf besar kecil. Dalam Data View, skor variabel sekarang akan muncul di sebelah kiri grup variabel.

MENGHAPUS DAN MENGATUR ULANG ITEM DARI TAMPILAN

Item dihapus dari tampilan dengan memilih dan menekan tombol **Delete**. Item dapat diatur ulang dengan mudah dengan mengklik dan

menyeretnya di panel kiri, panah merah yang menunjukkan tempat item akan dipindahkan saat Anda menyeretnya.

MEMBUAT JEDA HALAMAN (PAGE BREAK)

Jeda halaman di antara item yang termasuk kategori yang berbeda dapat dibentuk.

Langkah 1 Klik item di atas yang ingin Anda buat jeda halaman

Langkah 2 Pilih Insert dan kemudian klik Page Break

Langkah 3 Kembali ke tampilan dan klik di luar kotak pilihan untuk membatalkan pilihan. Periksa halaman di **Print Preview**.

MENGUBAH TIPE VARIABEL

Secara default, variabel type dalam kolom Type diasumsikan numerik. Untuk membuat **variabel string** Anda harus melanjutkan langkah-langkah sebagai berikut:

Langkah 1 Setelah mengetikkan nama variabel, sorot sel di kolom Type. (Lihat Langkah 1-3 yang diberikan di halaman sebelumnya).

Langkah 2 Klik area abu-abu dengan 3 titik di sebelah kanan Numeric untuk membuka Variable Type di kotak dialog.

Langkah 3 Klik tombol radio String dan kotak Width dan Decimal Places akan segera diganti dengan kotak berlabel Karakter.

Langkah 4 Ubah Value default 8 di kotak karakter ke nomor yang lebih besar seperti 20 untuk mengakomodasi nama terpanjang. Lakukan ini dengan memindahkan kursor ke dalam kotak angka memilih 8 dan kemudian menggantinya dengan mengetik 20.

Langkah 5 Klik OK. Sekarang kolom Width akan menampilkan 20. Pada Variable View, tipe variabel String akan muncul di bawah kolom Type.

Langkah 6 Klik kolom **Width column** dan salin spesifikasi baik dengan memilih **copy**.

Langkah 7 Klik **Column** dan tempel spesifikasi **Width** baru (20) dengan memilih **Paste**. Setelah ini, akan ada cukup ruang di Data View untuk melihat nama terpanjang di kumpulan data.

NILAI HILANG (MISSING)

SPSS mengasumsikan bahwa semua data set sudah lengkap. Namun, pengguna mungkin tidak memiliki entri untuk setiap kasus pada setiap variabel dalam kumpulan data. Entrian yang hilang tersebut ditandai oleh SPSS dengan apa yang dikenal sebagai Value **system-missing**, yang ditunjukkan dalam **Data Editor** dengan titik. SPSS akan mengecualikan system-missing dari perhitungannya, rata-rata hitung, standar deviasi dan statistik lainnya.

Berbeda dengan masalah di atas, terkadang pengguna mungkin ingin SPSS memperlakukan respons tertentu yang sebenarnya ada dalam kumpulan data sebagai data yang tidak tercantum (hilang) baik secara sengaja maupun tidak. Ini disebut **Missing Value**. Untuk menentukan Value yang hilang oleh pengguna tersebut, lakukan langkah-langkah berikut:

Langkah 1 Arahkan ke **Variable View**, gerakkan kursor ke kolom missing dan klik pada sel yang sesuai dari variabel yang bersangkutan.

Langkah 2 Klik area abu-abu dengan tanda pengganti di sebelah kanan **None** untuk membuka kotak dialog Missing Value.

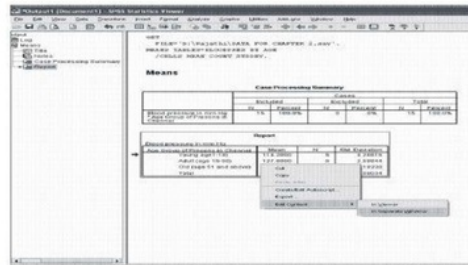
Langkah 3 Ada tiga tombol. Tombol **No Missing Value** ditandai sebagai **default**. Untuk variabel kuantitatif, klik rentang plus satu tombol missing yang terpisah. Masukkan Value 0 hingga 20 ke dalam kotak **Low** dan **High** masing-masing dan 9 ke dalam kotak Value Diskrit.

Langkah 4 Klik OK dan nilai-nilai akan muncul di sel kolom missing.

MENGEDIT OUTPUT SPSS

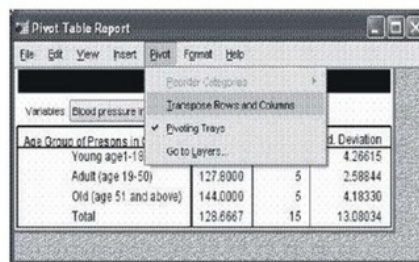
Jendela penampil Output SPSS dibagi menjadi dua panel oleh batang abu-abu vertikal (Lihat Gambar 1.18). Panel kiri menunjukkan organisasi hierarkis dari konten. Panel kanan menunjukkan hasil analisis statistik. Isi keduanya dapat diedit, karena tersedianya fasilitas pengeditan.

Langkah 1 Klik kanan pada tabel, batas lipatan muncul di atas dan pada menu pilih **Edit Contents** dan kemudian pilih **In a Separate Window** (Gambar 1.21), di mana variabel kategori muncul dalam baris dan deskripsi dalam kolom.



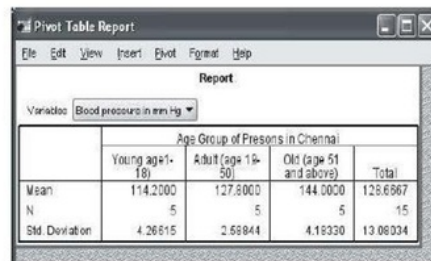
Gambar 1.21 Memilih opsi Edit Content dalam output SPSS

Langkah 2 SPSS Pivot Table Report muncul di atas. Anda dapat mengedit seperti yang Anda lakukan di Word. Jika Anda ingin statistik deskriptif dalam baris, kemudian klik pada Pivot dan kemudian pilih **Transpose Rows and Columns** (Gambar 1.22).



Gambar 1.22 Memilih Baris dan Kolom Transpose dalam Pivot Table Report

Langkah 3 Sekarang variabel kategori disalin dalam kolom dan deskripsi dalam baris (Gambar 1.23).



Gambar 1.23 Mengubah baris dan kolom judul setelah mengedit dalam output SPSS

MENYALIN OUTPUT SPSS

Prosedur optimal untuk menyalin item dalam tampilan sedikit berbeda untuk tabel dan grafik.

Menyalin Tabel

Langkah 1 Jika Anda berniat untuk menyalin tabel dari tampilan output di SPSS, simpan kursor di mana saja di dalam tabel dan klik, sekarang tabel tersebut berada dikotak.

Langkah 2 Klik Copy di menu Edit.

Langkah 3 Lanjut ke Word, pilih **Paste Special** dari Menu Edit dan kemudian pilih **Picture** dari jendela **Pop up**, atau jika Anda ingin menyalin lebih dari satu tabel, maka pastikan bahwa semua tabel yang diinginkan dikotak dengan menekan tombol **Ctrl** sambil mengklik setiap tabel secara bergantian dan kemudian Copy.

Pilih **Paste Special** di menu Edit, pilih Picture, lalu pilih Paste Special dari jendela Pop up.

Menyalin Grafik

Langkah 1 Tahan kursor di mana saja dalam grafik dan klik. Sekarang grafik memiliki kotak di sekitarnya. Jika lebih dari satu grafik akan disalin, tahan tombol Ctrl dan klik pada setiap grafik.

Langkah 2 Klik Copy di menu Edit.

Langkah 3 Beralih ke Word dan pastikan bahwa kursor berada di titik penyisipan.

Langkah 4 Klik Edit dan pilih Paste Special dan pilih Picture (Peningkatan Metafile). Item tersebut kemudian dapat dipusatkan, diperbesar atau diperkecil dengan mengkliknya sehingga memperoleh kotak di sekitarnya dengan tab biasa.

Mengubah dari tampilan Potret ke lanskap

Untuk mengubah dari Potret ke Lanskap ⁴⁷ ikuti langkah-langkah berikut ini:

Langkah 1 Klik tombol **Page Setup** di bagian atas kotak dialog **Print preview**. Lihat di **Orientation panel**, tombol diatur dalam **Portrait** (default). Ubah ke **Landscape**.

Langkah 2 Klik OK untuk kembali ke viewer.

Langkah 3 Akses kotak dialog **Print** dengan memilih **File** dan kemudian klik **Print**. Perhatikan pilihan **Print Range** di area kiri bawah kotak. Pilih opsi yang diinginkan, **All visible output** atau **Selection** yang terlihat dan klik OK. Jika Anda memilih yang pertama, seluruh konten dari tampilan akan dicetak keseluruhan. Dalam kasus terakhir, Anda hanya akan mendapatkan output yang dipilih.

MENCETAK DARI SPSS

Baik data dan output dapat dicetak dalam SPSS. Ada perbedaan antara output pencetakan dari SPSS Viewer dan pencetakan data dari Data Editor.

Mencetak Output dari Tampilan

Output SPSS sangat banyak dan mencetak hasil apa saja termasuk pencetakan materi yang tidak relevan juga. Oleh karena itu, seseorang harus memanfaatkan fasilitas pengeditan dari yang terlihat untuk menghapus semua materi yang tidak relevan. Seseorang dapat memilih item dan mencetaknya. Terdapat dua cara untuk memilih item, dengan mengklik ikon item di panel kiri tampilan, atau dengan mengklik item itu sendiri di panel kanan. Cara lain, kotak persegi panjang dengan batasan sendiri akan muncul di sekitar item. Kemudian pilih **Print Preview** untuk melihat jendela SPSS viewer (output yang dipilih), yang hanya akan menampilkan item yang dipilih. Kemudian kembali ke kotak dialog cetak, lihat bahwa tombol **Selection** di panel rentang Cetak diaktifkan dan kemudian klik OK. Hanya item yang dipilih yang akan dicetak. Untuk memilih lebih banyak item, klik yang pertama dan tekan tombol **Ctrl** dan klik item lain yang ingin Anda pilih. Sekarang pilih **Print Preview** untuk melihat item yang dipilih. Kembali ke kotak dialog **Print** dan berikan opsi pencetakan.

Menutup SPSS

SPSS ditutup dengan memilih **Exit** dari menu File. Jika data atau output belum disimpan, kotak dialog default akan muncul dengan pertanyaan. Simpan konten Data Editor ke **Untitled?** atau Menyimpan konten atau Output Viewer ke Output1? Klik Yes, No atau tombol Cancel. Anda diberi kesempatan terakhir untuk menyimpan konten Anda. Pilih file output yang dibutuhkan dan simpan. Kalau tidak, mungkin terlalu besar untuk ditampung.

Tutorial di SPSS

Paket SPSS memiliki tutorial tentang berbagai aspek sistem termasuk penggunaan tampilan dan manipulasi tabel pivot. Tutorial dapat diakses dengan memilih Help dan memilih Tutorial dan mengklik dua kali untuk membuka menu Tutorial. Tombol di sudut kanan bawah setiap halaman tutorial memungkinkan pengguna untuk melihat daftar item (diperbesar) dan menavigasi maju dan mundur melalui tutorial (panah kanan dan kiri).

Catatan: Software standar pada umumnya menyediakan fasilitas tutorial untuk membantu pembaca dalam mempelajari lebih dalam lagi bagaimana menjalankannya.

Mengimpor Data

Anda dapat mengimpor data ke SPSS dari platform lain seperti Microsoft, Excel, dan SPSS untuk Macintosh. Ini juga dapat membaca format file tetap dengan variabel yang direkam di lokasi kolom yang sama untuk setiap kasus. Juga dimungkinkan untuk mengekspor data dan output SPSS ke aplikasi lain seperti pengolah kata **Word** dan **spreadsheet**.

MENGIMPOR FILE EXCEL

Berikut adalah langkah-langkah untuk mengimpor file Excel yang disimpan dalam folder:

Langkah 1 Klik direktori file.

Langkah 2 Pilih jenis file dari kotak bernama sebagai File type: dan sorot Excel (*.xls).

Langkah 3 Klik file yang sesuai dari daftar file name yang muncul di panel dalam file name.

Langkah 4 Klik buka untuk mendapatkan kotak dialog **Opening File Option**. Pilih Nama Variabel untuk mentransfer nama variabel excel ke SPSS Data Editor.

Langkah 5 Jika muncul pesan kesalahan yang menyatakan bahwa SPSS tidak dapat memuat lembar kerja Excel, mungkin perlu untuk kembali ke Excel dan menyimpan ulang file dalam format versi Excel yang berbeda, untuk copy and paste kolom data langsung ke SPSS tampilan data.

Langkah 6 Klik OK untuk mentransfer file ke SPSS. Tampilan variabel akan mencantumkan nama variabel dan tipenya dan tampilan data akan menampilkan data yang ditransfer dan nama variabel. SPSS Viewer akan mencantumkan nama, tipe dan format variabel. File tersebut kemudian dapat disimpan sebagai file data SPSS.

Bab 2

STATISTIK DESKRIPTIF DENGAN SPSS

² Data yang dikumpulkan oleh peneliti dapat disajikan dalam bentuk tabel, diagram atau grafik. Selain itu, dimungkinkan untuk mendeskripsikan data berdasarkan ukuran numerik. Ketika data statistik digambarkan dalam ukuran numerik disebut statistik deskriptif. Ada beberapa langkah untuk mendeskripsikan satu set data. Mereka umumnya diklasifikasikan menjadi dua jenis: (i) ukuran tendensi sentral (ii) ukuran dispersi.

Ukuran-ukuran tendensi sentral atau, umumnya, rata-rata, menggambarkan tema sentral dari data dan meringkas karakteristik dari keseluruhan data. Karena nilai-nilai ini menemukan distribusi pada beberapa nilai variabel, kadang-kadang, disebut sebagai ukuran lokasi. Ukuran tendensi sentral yang paling umum dan berguna adalah rata-rata hitung. Ada langkah-langkah lain, yang memiliki penggunaan terbatas di berbagai bidang, yaitu: median, modus, rata-rata hitung geometrik, rata-rata hitung harmonik dan rata-rata tertimbang.

Ukuran-ukuran dispersi menggambarkan besarnya penyebaran nilai-nilai di sekitar ukuran tendensi sentral, yaitu, seberapa jauh atau seberapa dekat nilai-nilai itu dengan rata-rata. Simpangan baku adalah ukuran dispersi yang paling penting dan umum. Ukuran lain dari dispersi dengan penggunaan terbatas adalah kisaran (range, penyimpangan kuartil dan penyimpangan rata-rata.

Selain itu, ada langkah-langkah tertentu lainnya yang berguna dalam menggambarkan aspek data yang tidak diilustrasikan oleh ukuran tendensi sentral dan dispersi. Ini adalah ukuran kecondongan dan kurtosis. Kecondongan menggambarkan sifat simetri suatu distribusi dan kurtosis menggambarkan tingkat konsentrasi nilai di sekitar rata-rata hitung suatu distribusi.

Cara sederhana untuk mendeskripsikan data adalah dengan mencari ukuran tendensi sentral, dispersi, kecondongan dan kurtosis. Semua ukuran ini secara kolektif dikenal sebagai statistik deskriptif. Ketika pembaca mulai menggunakan SPSS, Anda seharusnya memiliki pengetahuan yang bagus tentang statistik. ¹² gaimanapun deskripsi singkat pada aspek-aspek teoritis dari rata-rata hitung, median, modus, standar deviasi, kecondongan dan kurtosis disajikan untuk memungkinkan pembaca untuk me-refresh sebelum menafsirkan hasil.

UKKURAN-UKURAN TENDENSI SENTRAL

7

Rata-rata Hitung

Rata-rata Hitung (Arithmetic Mean) didefinisikan sebagai jumlah dari semua item dari variabel dibagi dengan jumlah total item dalam sampel.

$$\text{Mean} = \frac{\text{Sum of all the items in a sample}}{\text{Total number of items in the sample}}$$

Definisi rata-rata hitung dinyatakan dalam persamaan aljabar sebagai

$$\bar{X} = \frac{\sum x}{n}$$

dimana,

\bar{X} – Rata-rata hitung variabel X

$\sum x$ – Jumlah semua item dari variabel

n – Jumlah total item dalam sampel-variabel X dari 1 hingga n.

Rumus di atas digunakan ketika nilai individual diberikan. Tetapi ketika data diklasifikasikan, rumus yang berbeda digunakan berdasarkan apakah seri tersebut diskrit atau kontinyu.

Seri diskrit

Dalam kasus seri diskrit, di mana frekuensi diberikan, rata-rata hitung aritmetik dapat dihitung dengan menerapkan rumus berikut

$$X = \frac{\sum f_i x_i}{\sum f_i}$$

dimana,

f –frekuensi

x – variabel

i – variasi x dari Value 1 hingga n.

Seri Kontinyu

Dalam kasus seri kontinyu, di mana frekuensi diberikan terhadap kelas interval, rata-rata hitung dapat dihitung dengan menerapkan rumus berikut

$$\bar{X} = \frac{\sum f_i x_i}{\sum f_i}$$

dimana,

f – frekuensi

m – titik tengah kelas interval variabel

$$\left(\text{Mid-point} = \frac{\text{lower limit} + \text{upper limit}}{2} \right)$$

MEDIAN

Median didefinisikan sebagai nilai tengah atau item dari kumpulan data yang diatur dalam urutan tertinggi atau terendah. Ini membagi seri menjadi dua bagian yang sama, satu bagian yang terdiri dari semua nilai lebih besar dari median dan bagian lain yang terdiri dari semua nilai kurang dari nilai median. Ini dianggap sebagai rata-rata posisi.

Seri tunggal

$$M = \frac{N+1}{2} \text{th item}$$

dimana,

N – jumlah total item

M – Median

Seri diskrit

$$M = \frac{N+1}{2} \text{th item}$$

dimana,

N – frekuensi total

Seri kontinyu

dimana,

L – batas bawah kelas median

H – besarnya frekuensi kelas median dari kelas median

F – adalah frekuensi total

C – Frekuensi kumulatif kelas yang mendahului (sebelum) kelas median

MODUS

Modus adalah titik konsentrasi maksimum. Ini adalah nilai dalam data yang diulang jumlah maksimum – didefinisikan sebagai nilai variabel yang paling sering terjadi dalam sampel.

Seri kontinyu

$$Z = l + \frac{f - f_1}{2f - f_1 - f_2} \times C$$

dimana,

Z – modus

l – batas bawah kelas modal

f – ³⁴ frekuensi kelas modal

f1 – frekuensi kelas yang sebelum kelas modus

f2 – frekuensi kelas setelah kelas modus

c – interval kelas distribusi frekuensi dapat menjadi unimodus.

Jika hanya satu nomor yang berulang kali jumlah maksimum, maka sampel memiliki modus tunggal. Jika sampel memiliki dua modus, itu disebut bimodus. Sampel multimodus atau polymodus juga terjadi.

UKURAN PENYEBARAN (DISPERSI)

Ukuran tendensi sentral salah satu aspek dari data yaitu. posisi sentralnya. Namun, ukuran ini saja tidak cukup untuk menggambarkan data sepenuhnya. Perbedaan antara nilai item dari rata-rata hitung (setiap ukuran tendensi sentral) disebut deviasi. Rata-rata penyimpangan nilai berbagai item dari ukuran tendensi sentral disebut ukuran dispersi.

STANDAR DEVIASI

Standar deviasi didefinisikan sebagai akar kuadrat rata-rata hitung dari penyimpangan kuadrat dari berbagai item rata-rata hitung. Singkatnya, ini disebut akar-rata-rata hitung-kuadrat-deviasi. Rata-rata hitung of square deviations disebut varians. Oleh karena itu, akar kuadrat dari varians adalah standar deviasi.

$$\sigma = \sqrt{\frac{\sum d^2}{N}}$$

Standar Deviasi,

di mana,

X - variabel

σ - standar deviasi

\bar{X} - rata-rata hitung variable $\bar{X} = \frac{\sum X_i}{n}$

d_i - deviasi $X - \bar{X}$

d_i^2 - deviasikuadrat $(X - \bar{X})^2 =$

Seri diskrit

$$\sigma = \sqrt{\frac{\sum fd^2}{N} - \frac{(\sum fd)^2}{N^2}}$$

dimana,

f –frekuensi

N – total

KECONDONGAN (SKEWNESS)

Ukuran tendensi sentral dan dispersi seperti rata-rata hitung dan standar deviasi masing-masing menggambarkan dua aspek penting dari distribusi data, nilai pusat dan tingkat penyebaran nilai-nilai di sekitar nilai pusat ini. Namun, ini tidak menggambarkan semua aspek distribusi. Misalnya, dua distribusi mungkin memiliki arti yang sama dan masih berbeda. Salah satu dari dua distribusi mungkin distribusi simetris dan yang lain mungkin asimetris. Distribusi asimetris dikatakan sebagai distribusi kecondongan, yaitu distribusi kecondongan. Kecondongan menggambarkan tingkat asimetri dalam suatu distribusi.

Distribusi condong positif

Distribusi yang condong positif memiliki sifat-sifat berikut:

- i. Lebih banyak item di kanan ordinat tertinggi (garis vertikal yang ditarik dari sumbu X ke kurva), yaitu, modus.
- ii. Rata-rata hitung > Median > Modus.
- iii. Kurva frekuensi memiliki kenaikan yang curam dan penurunan yang lambat dengan ekor panjang di sebelah kanan.

Distribusi condong negatif

Distribusi yang condong negatif memiliki properti berikut:

- i. Lebih banyak item di sebelah kiri ordinat tertinggi.
- ii. Rata-rata hitung < Median < Modus.
- iii. Kurva frekuensi memiliki kenaikan yang lambat dan jatuh dalam dengan ekor panjang di sebelah kiri.

Nilai yang berjarak sama dari modus tidak memiliki frekuensi yang sama. Ukuran numerik dikembangkan untuk mengevaluasi kecondongan suatu distribusi. Ini disebut koefisien kecondongan *Karl Pearson*.

$$\begin{aligned}\text{Karl Pearson's coefficient of skewness} &= \frac{(\text{Mean} - \text{Mode})}{\text{Standard deviation}} \\ &= -1 \text{ to } +1\end{aligned}$$

Ketika modus tidak terdefinisi dengan baik atau ketika distribusi memiliki lebih dari satu modus, maka

$$\begin{aligned}\text{Karl Pearson's coefficient of skewness} &= \frac{3 (\text{Mean} - \text{Median})}{\text{Standard deviation}} \\ &= -3 \text{ to } +3\end{aligned}$$

Jika nilainya negatif, distribusi condong negatif dan kurva frekuensi condong negatif dengan ekor panjang ke kiri. Jika positif, distribusi condong positif dan kurva frekuensi condong positif dengan ekor panjang ke arah kanan. Jika nilainya 0, distribusinya simetris. Nilai apa pun, + atau -, antara 0 dan 1 atau 0 dan 3, mencerminkan tingkat kecondongan.

KURTOSIS

Kedataran atau puncak dari kurva frekuensi digambarkan dengan ukuran yang disebut kurtosis. Distribusi frekuensi yang normal memberikan kurva berbentuk lonceng. Itu disebut *mesokurtic*. Kurva yang datar disebut *platykurtic*. Ketika frekuensi lebih atau kurang didistribusikan secara merata yaitu, frekuensi kurang terkonsentrasi di sekitar rata-rata memberikan kurva datar.

Ketika item terkonsentrasi lebih dekat dengan nilai rata-rata, yaitu, frekuensi lebih tinggi di tengah, itu memberikan kurva memuncak. Kurva memuncak disebut *leptokurtic*.

Kurtosis dari variabel acak adalah rasio momen sentral keempatnya terhadap kekuatan keempat standar deviasinya.

Koefisien Kurtosis

Jika koefisien kurtosis = 3, maka kurva frekuensi adalah mesokurtik; jika > 3, maka leptokurtik; jika < 3 maka itu adalah platykurtic.

STATISTIK DESKRIPTIF DENGAN SPSS

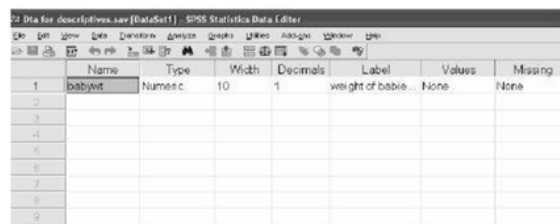
Contoh Data Kuantitatif 2.1

Berat bayi (kg) di bawah 6 bulan yang diambil dari catatan rumah sakit diberikan di bawah ini. Hitung rata-rata, median, modus, standar deviasi dan koefisien kecondongan dan kurtosis.

3.0	4.5	4.3	2.5	3.5	2.5	4.0	4.5	6.5	5.0
4.0	5.0	4.1	4.2	4.3	4.5	3.3	3.5	3.6	5.3
5.4	5.5	5.5	5.7	5.8	5.6	5.8	5.9	6.0	3.4
6.1	6.2	6.3	5.5	6.3	6.3	7.0	4.0	3.4	5.0

Langkah 1 Buka Data Editor SPSS.

Langkah 2 Klik Variabel View dan beri nama variabel sebagai "babywt". Pilih Jenis sebagai Numerik. Biarkan kolom standar Width tetap seperti itu, pilih "3" di bawah Desimal dan ketik "Weight of babies (kg) below six months" di bawah Label. Tidak perlu memberikan Value di bawah kolom Value karena datanya numerik (Gambar 3.1).



Data for descriptives.sav [DataSet1] - SPSS Statistics Data Editor							
	Name	Type	Width	Decimals	Label	Values	Missing
1	babywt	Numeric	10	1	weight of babies	None	None
2							
3							
4							
5							
6							
7							
8							
9							

Gambar 2.1 Penamaan variabel (babyweight) dalam Variabel View

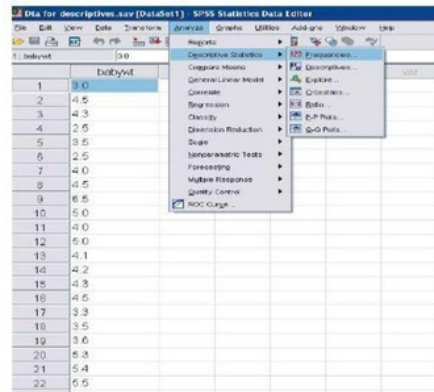
Langkah 3 Sekarang klik Data View dan letakkan Value di bawah kolom pertama di mana nama tersebut muncul seperti yang diberikan pada Gambar 2.2.

The screenshot shows the SPSS Data Editor window with a dataset named 'Data for descriptions.sav'. The variable 'babywt' is selected in the first column. The data is entered in the second column, with values ranging from 3.0 to 6.5. The first row is highlighted in blue.

1. Interval	babywt
1	3.0
2	4.5
3	4.3
4	2.5
5	3.5
6	2.5
7	4.0
8	4.5
9	6.5
10	5.0
11	4.0
12	5.0
13	4.1
14	4.2
15	4.3
16	4.5
17	3.3
18	3.5
19	3.6
20	5.3
21	5.4
22	6.5

Gambar 2.2 Memasukkan data (weight) di Data Editor

Langkah 4 Pilih Analisis dari menu utama, klik Statistik Deskriptif, lalu pilih Frekuensi. Menu akan muncul seperti Gambar 2.3.



Gambar 2.3 Memilih statistik deskriptif dari menu utama

Kotak dialog Frekuensi muncul seperti yang diberikan pada Gambar 2.4. Mentransfer variabel Berat bayi (kg) ke dalam Variabel Periksa tampilan Berat bayi (kg) di bawah Variabel.

Langkah 5 Klik OK untuk melanjutkan.



Gambar 2.4 Frekuensi kotak dialog dengan variabel yang ditransfer

Langkah 6 Klik Statistik untuk membuka Frekuensi: Statistik kotak dialog dan pilih Mean, Median dan Modus di bawah Tendensi Sentral. Standar Deviasi dan Varians di bawah Dispersi, Skewness dan Kurtosis di bawah Distribusi (Gambar 2.5).



Gambar 2.5 Frekuensi: Kotak dialog statistik dengan statistik deskriptif terpilih

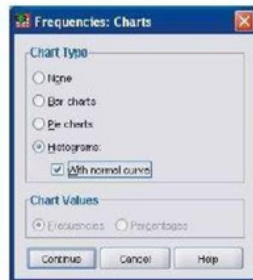
Langkah 7 Output statistik muncul di layar seperti di bawah ini (Output 1).

Output 1

Frequencies		
[DataSet1] C:\Documents and Settings\...		
Statistics		
Valid of output Statistics		
N	Valid	40
	Missing	0
Mean		4.925
Median		5.000
Mode		4.0 ^a
Std. Deviation		1.1558
Skewness		-.293
Std. Error of Skewness		.274
Kurtosis		-.956
Std. Error of Kurtosis		.733

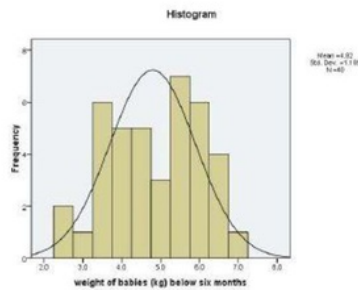
^a Multiple modes exist. The smallest value is shown.

Catatan: Anda dapat menggunakan prosedur yang sama baik untuk deskriptif maupun untuk grafik. Jika Anda ingin membuat histogram untuk data di atas, buka menu utama, pilih Descriptives dan kemudian Frekuensi untuk mendapatkan kotak dialog. Sekarang klik Charts... untuk membuka Frequencies: Charts. Dalam kotak Tipe Grafik, klik tombol Histogram dan tandai kurva normal dengan mengkliknya (Gambar 2.6). Klik Lanjutkan untuk kembali ke kotak dialog Frekuensi. Klik OK untuk menjalankan analisis. Output tampak seperti ditunjukkan pada Output 2.



Gambar 2.6 Frekuensi: Kotak dialog Charts dengan histogram dan kurva normal terpilih

Output 2



Histogram dengan kurva normal

Prosedur yang dijelaskan di atas berkaitan dengan data mentah, kadang-kadang ada kebutuhan untuk bekerja dengan data rahasia, dalam kasus seperti itu kita perlu mengikuti prosedur yang berbeda.

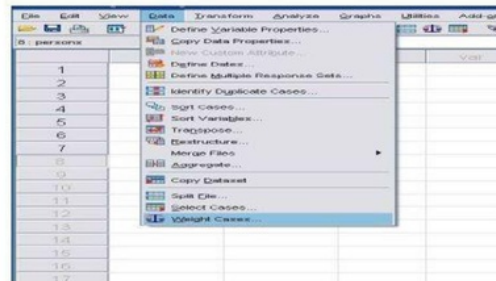
Contoh 2.2

Tabel berikut memberikan jumlah jam kerja dan jumlah orang untuk menyelesaikan tugas tertentu. Hitung rata-rata, median, modus, standar deviasi, skewness dan kurtosis.

Number of working hours	5	6	7	8	9	10	11
Number of persons	10	12	21	15	10	7	4

Langkah 1 Beri nama variabel dalam Variabel View dan masukkan data dalam Data View.

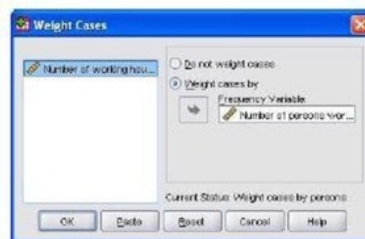
Langkah 2 Klik Data di menu utama, pilih Weight Cases (Gambar 2.7) (Langkah ini penting untuk data rahasia, karena memberikan bobot untuk frekuensi kelas).



Gambar 2.7 Memilih Weight cases dari menu utama

Langkah 3 Pilih Weight cases dengan tombol di bawah Weight cases.

Langkah 4 Transfer Jumlah orang (yaitu, frekuensi data) ke Frekuensi Variabel, lalu klik OK (Gambar 2.8). Sekarang tampilan menghilang dari layar.



Gambar 2.8 Kotak dialog Weight cases dengan jumlah orang yang ditransfer ke Variabel Frekuensi

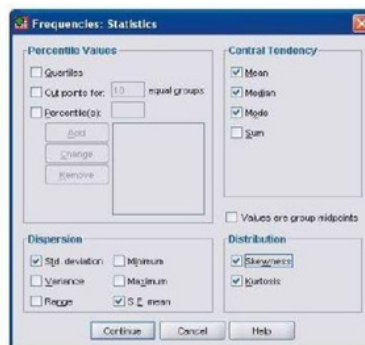
Langkah 5 Pilih Analisis dari menu utama. Klik Statistik Deskriptif dan pilih Frekuensi.

Langkah 6 Kotak dialog Frekuensi muncul. Transfer variabel Jumlah jam ke dalam Variable. Periksa tampilan Jumlah jam di bawah Variable (Gambar 2.9a).



Gambar 3.9a Frekuensi kotak dialog dengan Jumlah jam kerja ditransfer ke Variable

Langkah 7 Klik Statistik untuk membuka Frekuensi: Kotak dialog Statistik dan pilih rata-rata, median dan modus di bawah Tendensi Sentral. Simpangan baku dan SE Means di bawah Dispersi. Kecondongan dan kurtosis di bawah Distribusi (Gambar 2.9b).



Gambar 2.9b Frekuensi: Kotak dialog statistik dengan Descriptives dipilih di bawah judul berbeda

Langkah 8 Kemudian klik Lanjutkan dan kemudian klik OK untuk menjalankan analisis.

Langkah 9 Output muncul dengan statistik deskriptif seperti means, median dan modus, standar deviasi dan SE means, kecondongan dan kurtosis seperti yang ditunjukkan pada output.

Output

Statistics		
Number of working hours		
N	Valid	79
	Missing	0
Mean		7.5063
Std. Error of Mean		.18625
Median		7.0000
Mode		7.00
Std. Deviation		1.65540
Skewness		0.33
Std. Error of Skewness		0.27
Kurtosis		-0.58
Std. Error of Kurtosis		0.53

Tabel statistik deskriptif

Catatan: Prosedur serupa diadopsi untuk dikerjakan dalam seri kontinyu.

DATA NOMINAL DAN ORDINAL ATAU DATA KUALITATIF

Frekuensi dari variabel nominal atau kategori diwakili dalam baris dan kolom dalam bentuk tabel. Ini bisa diperoleh dengan beberapa cara. Tabel frekuensi (dari menu Tabel) dan Crosstabs (dari Statistik Deskriptif) menyediakan tabel kontingensi dua arah. Selain itu Crosstabs dari Statistik Deskriptif juga memberikan frekuensi dengan total baris dan kolom (prosedur ini juga menyediakan statistik seperti Chi-square dan koefisien Korelasi). Frekuensi dalam Statistik Deskriptif memberikan distribusi frekuensi untuk variabel nominal dan ordinal. Ini juga memberikan persentase dan frekuensi kumulatif. Ada juga pilihan untuk memilih grafik seperti diagram batang, diagram lingkaran, dan histogram.

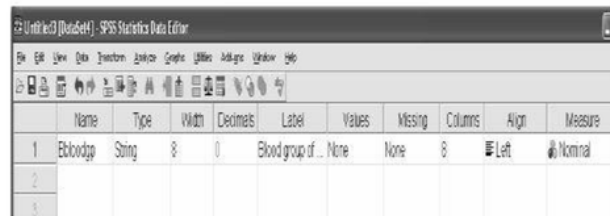
Data nominal: merumuskan tabel frekuensi

Contoh 2.3

Merumuskan tabel frekuensi dan menggambar diagram pie untuk data berikut pada golongan darah dari 45 mahasiswa.

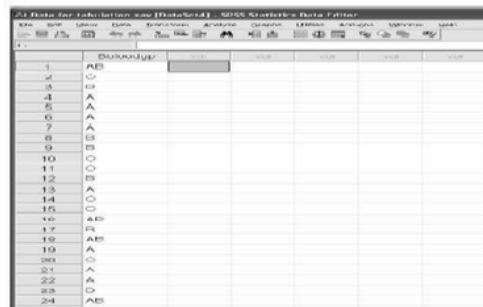
AB	B	O	A	O	O	A
O	B	AB	B	A	B	A
B	O	B	AB	A	O	O
A	O	AB	O	O	A	
A	B	A	A	AB	O	
A	A	O	A	O	A	
A	O	A	O	O	B	

Langkah 1 Beri nama variabel sebagai bloodgp dan klik pada sel bernama Type select String, label di bawah label Column sebagai Golongan Darah mahasiswa (Gambar 2.10).



Gambar 2.10 Penamaan grup darah variabel di bawah Variable View

Langkah 2 Masukkan kotak darah berdasarkan kasus di Data Editor seperti pada Gambar 2.11.



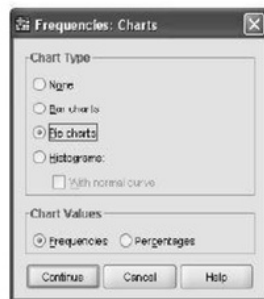
Gambar 2.11 Data pada golongan darah yang dimasukkan dalam Tampilan Data

Langkah-langkah berikut menawarkan frekuensi dan grafik.

Langkah 3 Pilih Analisis dari menu utama, pilih Statistik Deskriptif dan kemudian pilih Frekuensi untuk membuka kotak dialog Frekuensi (Gambar 2.12a) dan transfer golongan darah ke kotak variabel. Klik Charts untuk membuka Frekuensi: Charts dialog box dan pilih opsi Pie chart dan klik continue (Gambar 2.12b).



Gambar 2.12a Frekuensi kotak dialog dengan variabel yang ditransfer



Gambar 2.12b Frekuensi: bagan kotak dialog dengan opsi dipilih diagram Pie

Langkah 4 Klik Charts untuk mendapatkan Frekuensi: Charts kotak dialog dan pilih Chart. Ada juga opsi Chart Values. Jika Anda ingin menampilkan Value pada grafik, pilih Frekuensi atau Persentase di bawah Charts Value.

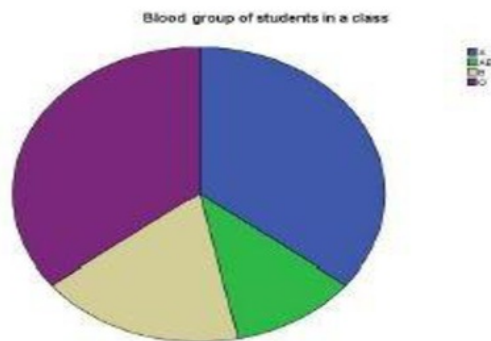
Langkah 5 Output muncul seperti ditunjukkan di bawah ini (Output 1).

Output 1

Blood group of students in a class				
		Frequency	Percent	Cumulative Percent
Valid	A	16	35.6	35.6
	AB	5	11.1	46.7
	B	8	17.8	64.4
	O	16	35.6	100.0
	Total	45	100.0	

Tabel frekuensi yang menunjukkan golongan darah mahasiswa di kelas

Output 2



Pie chart

Catatan: Jika Anda memilih Bar Charts di bawah Frekuensi kemudian barchart muncul di output seperti pada Output 3.

Output 3

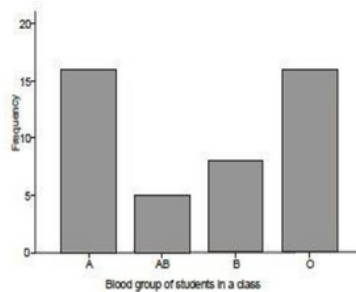


Chart batang sederhana

Data kategori: merumuskan tabel frekuensi

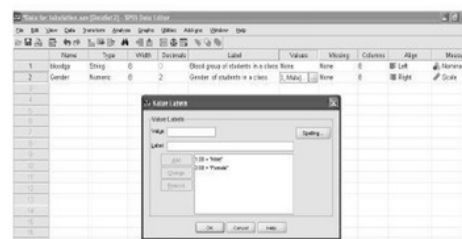
Contoh 2.4

Rumuskan tabel frekuensi dan gambarlah bar chart clustered (grafik batang bertumpuk) dan terbagi untuk data berikut pada kelompok darah 90 mahasiswa.

Male										Female									
AB	B	O	A	A	O	A	O	B	O	A	O	O	O						
O	B	AB	O	B	B	A	B	B	AB	B	A	B	B						
B	O	B	A	AB	O	O	O	O	B	AB	A	O	O						
A	O	AB	O	O	A				A	O	AB	O	O	A					
A	B	A	A	AB	O				O	B	A	A	AB	O					
A	A	O	A	O	A				A	A	O	A	O	A					
A	O	A	O	O	B				A	O	A	O	O	B					

Langkah 1 Beri nama variabel sebagai bloodgp dan klik pada sel bernama Type, pilih String. Ketikkan di bawah Label kolom sebagai "Blood Group" mahasiswa.

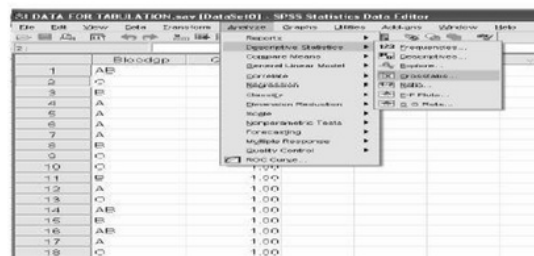
Langkah 2 Ketik variabel kedua di bawah Nama sebagai "Gender", di kolom kedua Type, tetap Numeric. Biarkan Width tetap sebagai default, kolom Desimal bisa menjadi 0. Ketik di bawah Label sebagai "Gender of students in a class" seperti pada Gambar 2.13.



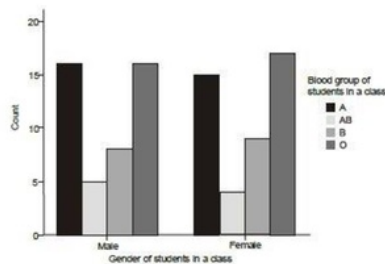
Gambar 3.13 Variabel view dengan Label Value dan Value berlabel

Langkah 3 Masukkan data dalam Data view.

Langkah 4 Pilih Analisis kemudian klik Statistik Deskriptif dan kemudian pilih Frekuensi... untuk membuka kotak dialog Frekuensi dan kemudian pilih Crosstabs (Gambar 2.14).



Gambar 2.14 Memilih opsi Crosstabs dari menu utama



Bar chart cluster yang menunjukkan golongan darah mahasiswa berdasarkan gender

REVIEW LATIHAN

1. Data panjang daun (cm) untuk dua kelompok pohon diberikan di bawah ini. Hitung mean, median, modus, standar deviasi, standard error, skewness dan kurtosis. Tafsirkan hasil Anda.

Group I	50	65	63	50	54	60	66	68	71	64
Group II	62	72	68	58	65	70	72	70	71	70

2. Hitung rata-rata hitung, standar deviasi, skewness dan kurtosis untuk data berikut pada ketinggian (cm) dari 40 mahasiswa dan uraikan distribusi.

151	152	156	175	151	164	174
185	168	160	155	153	165	163
170	167	165	157	164	167	170
174	163	158	165	168	183	172
155	180	159	169	171	155	
164	175	174	159	172	177	
185	168	160	155	153	165	

3. Hitung ¹² rata-rata hitung, median, modus, standar deviasi, standard error, skewness dan kurtosis untuk distribusi berikut.

Number of flowers (x)	1	2	3	4	5	6	7
Number of plants (y)	8	4	12	9	2	2	1

4. Status pernikahan pria di 2 tempat yang berbeda di kota tertentu diberikan di bawah ini. Merumuskan tabel frekuensi.

Street 1	M, M, M, S, M, M, S, M, S, M, M, M, M, S, M, M, M, S,
	M, S, M, M, M, M, S, M, M, M, M, M, M, M, M, M, M,
	S, M, M, M, M, M, M, M, M, M, M, M, S, S, S, M, M, M, M,
	M, M, M, M
Street 2	M, M, M, M, M, M, M, M, S, M, M, S, M, S, M, S, M, M, M,
	M, M, M, S, M, M, S, M, M, M, M, M, M, M, M, S, M, M,
	S, M, M, M, M, M, M, M, M, M, M, M, S, S, M, M, M, M

5. Kelompok darah mahasiswa di kelas diberikan di bawah ini. Buatlah tabel frekuensi.

Day scholars							Hostellers						
AB	B	O	A	A	O	A	O	B	O	A	O	O	
O	B	AB	O	B	B	A	B	B	AB	B	A	B	
B	O	B	A	AB	O	O	O	O	B	AB	A	O	
A	O	AB	O	O	A	A	A	O	AB	O	O	A	
A	B	A	A	AB	O	A	O	B	A	A	AB	O	
A	A	O	A	O	A	O	A	A	O	A	O	A	
A	O	A	O	O	B	B	A	O	A	O	O	B	

Bab 3

GRAFIK DAN DIAGRAM

SPSS menyediakan berbagai opsi untuk grafik dan diagram/bagan. Ada pilihan untuk Grafik di menu utama, dan ada juga opsi yang sama dalam prosedur analisis. Sebagai contoh ada pilihan Charts dalam prosedur Frequency dan opsi **Profil Plot** dalam prosedur ANOVA. Grafik dapat dibuat agar terlihat menarik dengan menggunakan **Chart Builder** atau Interaktif dari sub menu. Dengan mengklik pada item Chart Builder di menu Graph, galeri diagram dan grafik dapat diperoleh. Galeri berguna dalam memilih grafik atau bagan yang diinginkan.

GRAFIK BATANG

Grafik Batang ditujukan untuk variabel diskrit, baik kualitatif maupun kuantitatif. Ada bar sederhana grafik batang sederhana, diagram batang bergerombol dan grafik batang bertumpuk. Bagan batang adalah diagram satu dimensi di mana ketinggian bar sama dengan frekuensi atau rata-rata variabel. Satu memiliki pilihan untuk memilih salah satu tergantung pada kesesuaian data. Untuk variabel kuantitatif, diagram batang dengan bar kesalahan dapat ditarik, sedangkan untuk variabel enumerasi atau variabel kualitatif hanya diagram batang dapat digambar.

GRAFIK BATANG SEDERHANA

Diagram batang sederhana diinginkan untuk membandingkan rata-rata kelompok pengamatan (tinggi mahasiswa) atau frekuensi sederhana dari variabel kualitatif (golongan darah siswa di kelas).

Grafik batang sederhana untuk variabel kualitatif

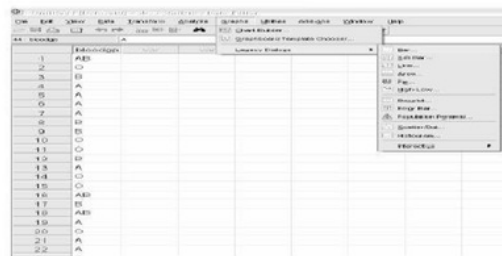
Contoh 3.1

Buatlah tabel frekuensi dan gambar diagram batang untuk data berikut pada kelompok darah dari 45 siswa di kelas.

AB	B	O	A	O	O	A
O	B	AB	B	A	B	A
B	O	B	AB	A	O	O
A	O	AB	O	O	A	
A	B	A	A	AB	O	
A	A	O	A	O	A	
A	O	A	O	O	B	

Langkah 1 Beri nama variabel dalam Variabel view dan masukkan data dalam data editor.

Langkah 2 Pilih Grafik dari menu utama dan klik Bar (Gambar 3.1) untuk membuka kotak dialog Bar chart (Gambar 3.2).



Gambar 3.1 Memilih opsi Bar chart dari menu utama



Gambar 3.2 Bar Charts untuk memilih Simple Bar

Langkah 23 Pilih Simple dan klik Define untuk membuka kotak centang Define Simple Bar: Summaries for Groups of Cases (Gambar 3.3).



1 Gambar 3.3 Define Simple Bar: Summaries for Group of Cases dengan golongan darah mahasiswa yang dipilih dalam sumbu kategori

Langkah 4 Transfer nama variabel ke sumbu kategori.

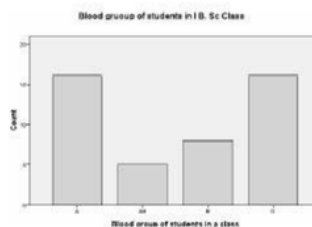
Langkah 5 Klik Judul untuk membuka kotak dialog dan masukkan judul dalam kotak teks (Gambar 3.4). Klik Lanjutkan untuk kembali ke kotak dialog diagram Batang dan kemudian klik OK untuk mendapatkan bagan.



Gambar 3.4 Judul kotak dialog dengan judul diketik di Baris 1 sebagai "Kelompok darah mahasiswa"

Langkah 6 Grafik muncul dalam output.

Output



Grafik Bar Sederhana untuk golongan darah mahasiswa

Grafik batang sederhana untuk variabel kuantitatif. Grafik batang cocok untuk variabel kuantitatif diskrit di mana ketinggian batang mewakili rata-rata sekelompok pengamatan dan garis vertikal dengan batang horizontal yang dipasang di atas batang mewakili standar kesalahan dari rata-rata hitung.

Oleh karena itu, dalam diagram batang dengan *bar error*, baik rata-rata hitung dan standar error diwakili.

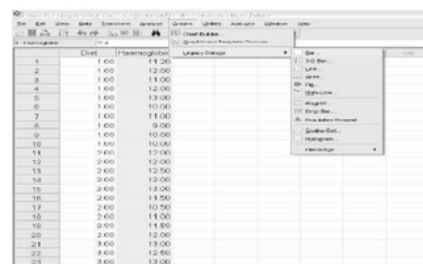
Contoh 3.2

Empat kelompok anak diberi makan empat diet yang berbeda dan tingkat hemoglobin mereka dihitung. Data diberikan dalam tabel berikut. Buatlah data dalam bentuk diagram batang dengan standar error.

Diet 1	11.2, 12.0, 11.0, 12.0, 13.0, 10.0, 11.0, 9.0, 10.0, 10.0
Diet 2	12.0, 12.0, 12.5, 12.0, 13.0, 11.5, 10.5, 11.0, 11.5, 12.0
Diet 3	13.0, 12.5, 13.0, 13.0, 12.0, 12.5, 11.0, 12.0, 11.5, 13.0
Diet 4	9.5, 9.6, 10.0, 12.0, 11.0, 9.8, 10.0, 10.0, 9.8, 9.9

Langkah 1 Beri nama variabel (sebagai kategorikal) di Variabel view dan masukkan data di Data Editor.

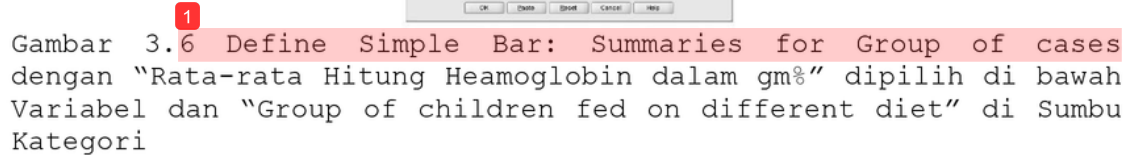
Langkah 2 Pilih Grafik dari menu utama kemudian klik **Dialog Legacy** dan kemudian pilih **Bar...** untuk membuka kotak dialog Bar charts (Gambar 3.5).



Gambar 3.5 Memilih opsi Bar Chart dari menu utama

Langkah 3 Pilih **Simple** dan klik **Define** untuk membuka kotak centang **Define Simple Bar: Summaries for Groups of Cases**.

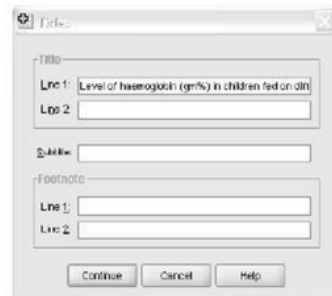
Langkah 5 Pilih tombol Statistik lainnya (mis. Rata-rata hitung) di bawah Bars Represent. Langkah ini penting untuk menampilkan rata-rata, dalam diagram batang, jika tidak, jumlah kasus akan diwakili dalam bagan batang.



Bar Sederhana.



Langkah 7 Klik **Titles** untuk membuka kotak dialog Titles (Gambar 3.8) dan masukkan Titles di kotak teks.



Gambar 3.8 **Titles** kotak dialog "Level of haemoglobin (gm%) in children fed on different diet"

Langkah 8 Klik Lanjutkan untuk kembali ke kotak dialog Bar chart dan kemudian klik OK untuk mendapatkan grafik. Output meringkas nilai rata-rata dengan standar error pada tingkat hemoglobin di antara anak-anak yang diberi diet yang berbeda muncul seperti yang diberikan dalam output.

Output

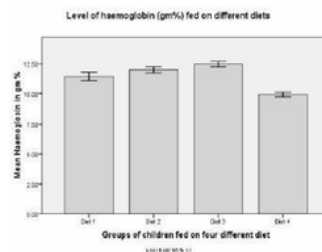


Diagram batang sederhana dengan bar kesalahan

GRAFIK BATANG KLASSTER

Diagram ini menunjukkan dua atau lebih kategori variabel dalam grafik yang sama. Diagram klaster dapat diplot untuk data berikut pada golongan darah dan jenis kelamin. Variabel kedua menunjukkan klasternya.

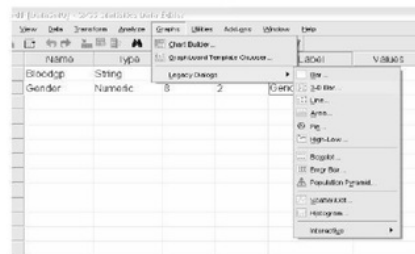
Contoh 3.3

Gambarlah diagram batang berlapis dan ditumpuk untuk data pada kelompok darah 90 mahasiswa.

Male						Female					
AB	B	O	A	A	O	A	O	B	O	A	O
O	B	AB	O	B	B	A	B	B	AB	B	A
B	O	B	A	AB	O	O	O	O	B	AB	A
A	O	AB	O	O	A		A	O	AB	O	O
A	B	A	A	AB	O		O	B	A	A	AB
A	A	O	A	O	A		A	A	O	A	O
A	O	A	O	O	B		A	O	A	O	B

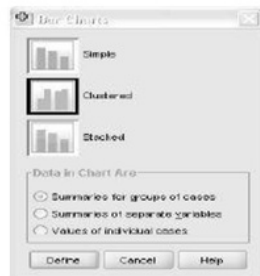
Langkah 1 Beri nama variabel dalam **Variabel View**. Masukkan data dalam **Data editor**.

Langkah 2 Pilih Grafik dari menu utama klik **Legacy Dialog** dan kemudian klik Bar untuk membuka Bar chart (Gambar 3.9).



Gambar 3.9 Memilih opsi Bar chart dari menu utama

Langkah 3 Pilih **Clustered** dan klik **Define** untuk membuka **Define Simple Bar:Summaries for Groups of Cases** (Gambar 3.10).



Gambar 3.10 Kotak dialog grafik batang untuk memilih grafik batang kluster

Langkah 4 Transfer nama variabel yaitu "Gender" ke Kategori Sumbu dan "Blood Group" untuk Definisikan Cluster oleh (Gambar

3.11). Untuk menjumlahkan data jenis ini pilih tombol **N of cases** pada **Bars Represent**.



Gambar 3.11 **Define Clustered Bar: Summaries for Group of Cases** untuk mentransfer bloodgp ke **Define Cluster by** dan gender ke sumbu Category.

Langkah 5 Klik Option untuk membuka kotak dialog Opsi dan klik Titles untuk membuka Judul kotak dialog (Gambar 3.12) dan masukkan judul di baris 1 di kotak teks.

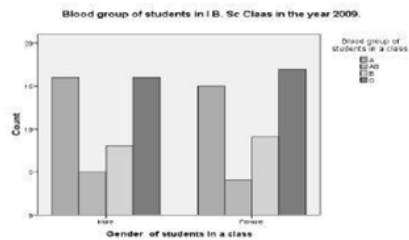


Gambar 3.12 Kotak dialog **Titles** dengan judul diketik di baris 1 sebagai "Blood group of student"

Langkah 6 Klik Lanjutkan untuk kembali ke kotak dialog Bar chart dan kemudian klik OK untuk mendapatkan grafik.

Langkah 7 Grafik muncul seperti pada output (Output 1).

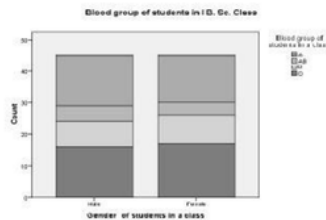
Output 1



Bar chart klaster

Jika Anda memilih **stacked** di kotak dialog diagram batang (Gambar 3.10) diagram batang bertumpuk muncul sebagai **Output 2**.

Output 2



Clustered bar chart with error bars

Diagram batang klaster atau diagram batang yang dipanel dari data set yang terdiri dari variabel yang diukur untuk 2 atau 3 faktor dengan masing-masing tingkat yang berbeda. Dalam contoh 3.2, tingkat hemoglobin pada anak-anak yang diberi empat diet yang berbeda diberikan. Empat jenis diet yang berbeda adalah tingkatan faktor yaitu diet. Jika data tersebut dikumpulkan dari tiga Puskesmas (PHC) maka puskesmas tersebut menjadi variabel kedua dan ketiga Puskesmas menjadi tingkatannya. Kita dapat merepresentasikan data semacam ini dengan diagram batang klaster atau dipanel dengan **bar charts with error bars**.

Contoh 3.4

Gambarlah diagram batang berpanel atau berlapis dengan bar kesalahan untuk data serupa yang diberikan dalam Contoh 3.2, untuk anak-anak yang menjadi anggota dua pusat kesehatan primer.

Langkah 1 Masukkan data seperti pada contoh 3.2. Sertakan variabel Puskesmas dalam Variabel view, beri Label pada variabel dan beri Value dengan pengkodean 1 sebagai PHC1, 2 sebagai PHC2 dan 3 sebagai PHC3 (Gambar 3.13).

	Hemoglobin	Puskesmas	Kelompok anak-anak
1	1.000	1.000	1.000
2	1.000	1.000	1.000
3	1.000	1.000	1.000
4	1.000	1.000	1.000
5	1.000	1.000	1.000
6	1.000	1.000	1.000
7	1.000	1.000	1.000
8	1.000	1.000	1.000
9	1.000	1.000	1.000
10	1.000	1.000	1.000
11	1.000	1.000	1.000
12	1.000	1.000	1.000
13	1.000	1.000	1.000
14	1.000	1.000	1.000
15	1.000	1.000	1.000
16	1.000	1.000	1.000
17	1.000	1.000	1.000
18	1.000	1.000	1.000
19	1.000	1.000	1.000
20	1.000	1.000	1.000
21	1.000	1.000	1.000
22	1.000	1.000	1.000
23	1.000	1.000	1.000

Gambar 3.13 Data Editor dengan 23 kasus (sebagian) dimasukkan untuk tiga variabel yang berbeda

Langkah 2 Ikuti langkah-langkah seperti yang diberikan untuk diagram batang bertumpuk dan transfer variabel yaitu, "Hemoglobin" ke dalam Variabel, "Puskesmas" ke dalam Sumbu Kategori dan "Kelompok anak-anak" ke dalam **Define Cluster by** (Gambar 3.14).



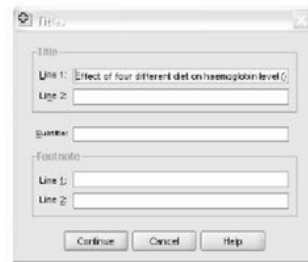
Gambar 3.14 Tentukan **Define Clustered Bar: Summaries for Group of Cases** dengan "MEAN Hemoglobin dalam gm%" dipilih untuk Variabel "Puskesmas" di sumbu Kategori dan "Kelompok anak yang diberi diet berbeda" dikotak "Define Cluster by".

Langkah 3 Klik Options untuk membuka kotak dialog Options, pilih Display error bars sebagaimana terlihat pada Gambar 3.15. Klik Lanjutkan untuk kembali ke Define Simple Bar.



Gambar 3.15 Kotak dialog **Options** dengan pilihan **Display error bars**

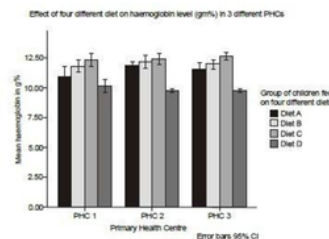
Langkah 4 Klik Judul untuk membuka kotak dialog Titles dan masukkan Judul di kotak teks (Gambar 3.16).



Gambar 3.16 Kotak dialog Titles dengan judul "Pengaruh empat diet berbeda pada hemoglobin (gm%) pada anak-anak dalam 3 PHC"

Langkah 5 Klik Lanjutkan untuk kembali ke kotak dialog Bar chart dan kemudian klik OK untuk mendapatkan grafik. Output memberikan nilai rata-rata dengan standar error tingkat hemoglobin di antara anak-anak yang diberi diet yang berbeda muncul dalam klater untuk tiga PHC yang berbeda seperti terdapat pada output.

Output



Bar chart klaster dengan error bars

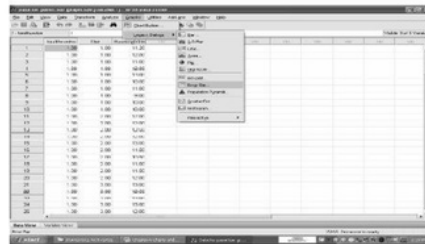
GRAFIK BATANG ERROR

Grafik batang error adalah alternatif grafik batang, di mana rata-rata sampel diwakili oleh satu titik dan penyebaran variabel, standar error diwakili oleh garis vertikal atau seperti huruf T yang melewati titik .

Contoh 3.5

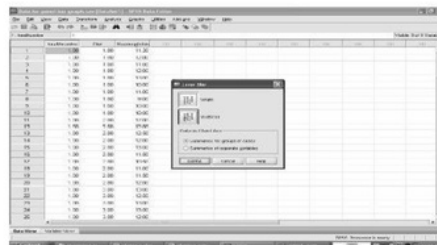
Gambarkan diagram error bars untuk data yang diberikan dalam Contoh 3.4.

Langkah 1 Pilih Grafik dari menu utama Legacy Dialog dan kemudian pilih Error bars ... untuk membuka kotak dialog Error Bars (Gambar 3.17).



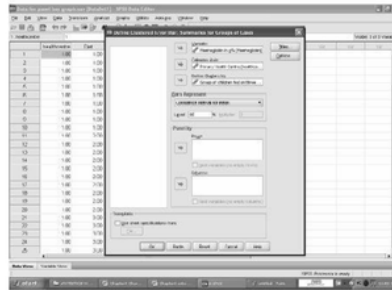
Gambar 3.17 Memilih Error bars dari menu utama

Langkah 2 Pilih Clustered dari Error Bar (Gambar 3.18).



Gambar 3.18 Kotak dialog Error Bars dengan Clustered dipilih

Langkah 3 Klik Define untuk membuka **Define Error Culstere⁵⁹ Bars: Summaries Group of Cases** dan variabel yang dipindahkan seperti terlihat pada Gambar 3.19.



Gambar 3.19 Define Clustered Error Bars: Summaries for groups of cases

Langkah 4 Klik Titles dan ketikkan judul seperti pada Contoh 3.4.

Langkah 5 Klik Lanjutkan dan kemudian klik OK.

Langkah 6 Output meringkas nilai rata-rata hitung tingkat hemoglobin dengan standar error sebagai huruf T di kedua sisi untuk anak-anak yang diberi empat diet yang berbeda muncul dalam kelompok untuk tiga PHC yang berbeda seperti pada output.

Output

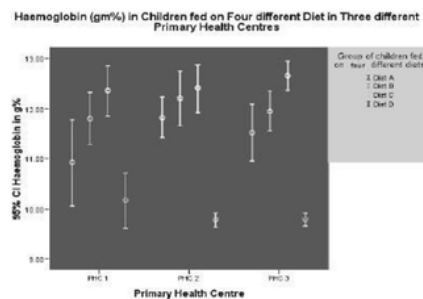


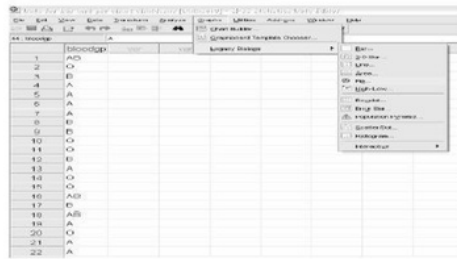
CHART "PIE"

Ini adalah diagram lingkaran di mana berbagai frekuensi kelas yang berbeda sama dengan sudut berbagai sektor yang berbeda di lingkaran. Digunakan untuk menampilkan frekuensi relatif dari kumpulan dataset yang sama. Frekuensi variabel kualitatif diwakili diagram Pie sebagai alternatif grafik batang. Data pada golongan darah dalam Contoh 3.1 dapat dibuat diagram lingkaran.

Contoh 3.6

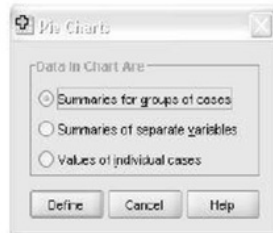
Buatlah Grafik Pie untuk data dalam Contoh 3.1.

Langkah 1 Pilih Graphs dari menu utama, klik Legacy Dialogs untuk membuka Pie charts (Gambar 3.20).



Gambar 3.20 Memilih opsi Pie Chart dari menu utama

Langkah 2 Klik Define kotak dialog Define Pie: Summaries for Groups of Cases (Gambar 3.21).



Gambar 3.21 Grafik Pie untuk memilih **Summaries for Groups of Cases**

Langkah 3 Klik variabel (golongan darah) dan pindahkan ke dalam **Define Slice by**. Jika Anda inginkan dalam %, pilih % of cases, jika tidak N of cases (Gambar 3.22).



Gambar 3.22 **Define Pie:Summaries for Groups of Cases** dengan "Kelompok darah siswa" dipilih dalam **Define Slice by**

Langkah 4 Akhirnya klik Titles untuk membuka kotak dialog Titles. Ketik judul, lalu klik Lanjutkan untuk kembali ke kotak Define Pie.

Langkah 5 Klik OK untuk mendapatkan diagram lingkaran. Output yang menyimpulkan kelompok darah di sektor-sektor lingkaran seperti tampak pada Output 1.

Output 1

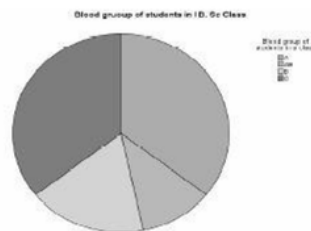


Diagram lingkaran menunjukkan Kelompok Darah mahasiswa

Langkah 6 Dari output, setiap irisan dapat diberi label di dalam atau diluar lingkaran, jika diperlukan untuk menekankan kategori tertentu. Ini dilakukan dengan menyorot slice dan memilih **Explode Slice**. Output yang menyajikan irisan dengan penekanan untuk golongan darah AB muncul seperti pada Output 2.

Output 2

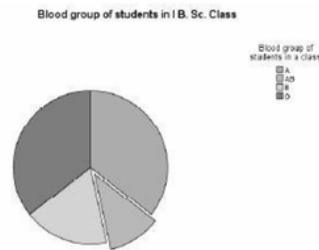


Diagram lingkaran yang menunjukkan golongan darah dengan tekanan irisan golongan darah AB

DIAGRAM PENCAR DAN DIAGRAM TITIK

Scatter Plots (Diagram Pencar)

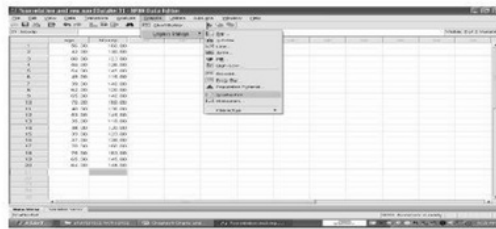
Hubungan antara dua variabel kuantitatif dapat dibuat dalam bentuk scatter plot. Skala nilai dari dua variabel seperti usia dan tekanan darah sistoliknya ditetapkan pada sumbu horizontal (sumbu X) dan vertikal (sumbu Y) masing-masing, setiap orang diwakili oleh sebuah titik dan setiap titik pada gilirannya mewakili usia dan tekanan darahnya. Diagram pencar tersebut harus selalu diplot dan diteliti sebelum melakukan perhitungan koefisien korelasi atau analisis regresi.

Contoh 3.7

Gambar diagram pencar untuk data berikut tentang usia (tahun) versus tekanan darah sistolik (mmHg.)

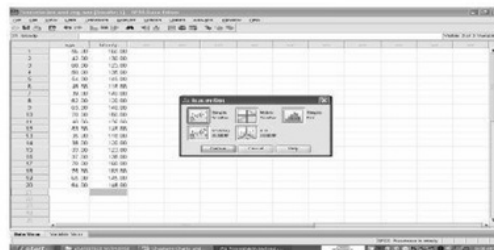
Age	56	42	60	50	54	49	39	62	65	70	40	53	35	38	39	37	70	75	65	64
BP	160	130	125	135	145	115	140	120	140	160	126	145	118	120	123	138	160	163	145	146

Langkah 1 Pilih Grafik dari menu utama untuk membuka kotak dialog **Scatter/Dot** (Gambar 3.23).



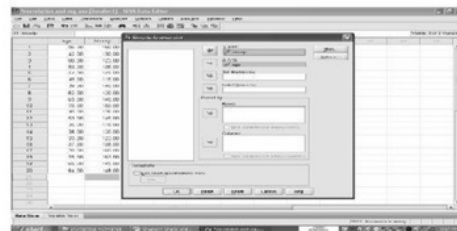
Gambar 3.23 Memilih Scatter/Dot dari menu utama

Langkah 2 Klik Define untuk membuka kotak dialog **Simple Scatter** (Gambar 3.24).



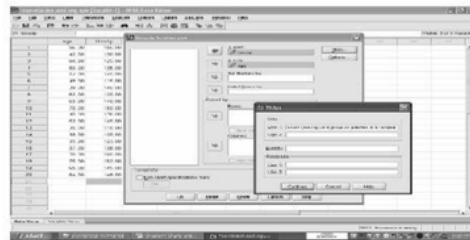
Gambar 3.24 Scatter Dot dengan pilihan Simple Scatter

Langkah 3 Pindahkan nama variabel seperti pada Gambar 3.25.



Gambar 3.25 Simple Scatter dengan variabel tekanan darah dan usia yang dipilih masing-masing pada sumbu y dan x

Langka²⁰ 4 Klik Titles untuk mendapatkan kotak Titles dan ketik judul seperti yang ditunjukkan pada Gambar 3.26.

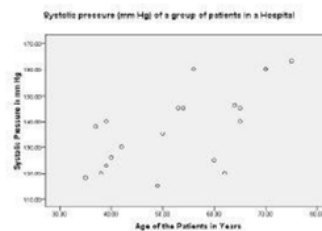


Gambar 3.26 Kotak Titles dengan judul yang diketik di Baris 1

Langkah 5 Klik Lanjutkan dan kemudian OK untuk mendapatkan diagram pencar.

Langkah 6 Output muncul dengan "Umur pasien dalam sumbu-x dan tekanan Sistolik dalam mmHg di sumbu -y dengan titik-titik untuk setiap orang seperti pada output.

Output



Sebarikan plot atau Dot plot

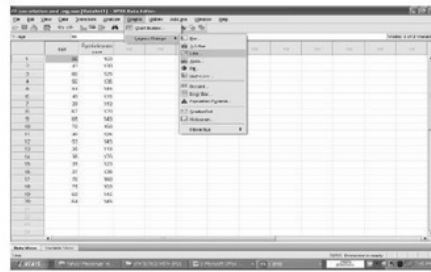
GRAFIK GARIS (LINE GRAPHS)

Grafik garis digambar untuk dua variabel atau lebih dari dua variabel. Grafik garis dapat digambar hanya dengan satu baris atau lebih dari satu baris dalam grafik. Contoh di atas pada tekanan darah (mmHg) individu dalam hubungannya dengan usia dapat dibuat dalam bentuk grafik garis; kisaran total tekanan darah dibagi menjadi interval tetap.

Contoh 3.8

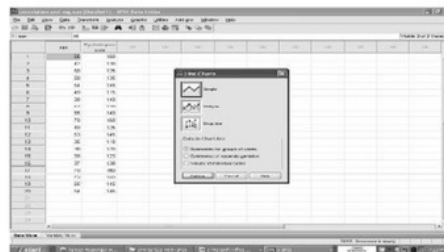
Gambarkan grafik garis untuk usia versus tekanan darah sistolik untuk data dalam Contoh 3.6.

Langkah 1 Masukkan data, pilih Graphs dari menu utama, pilih Legacy Dialog dan klik **Line** untuk membuka kotak dialog **Line Chart** (Gambar 3.27).



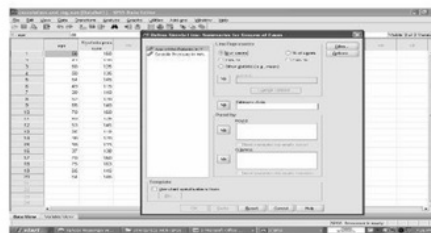
Gambar 3.27 Memilih opsi Line dari menu utama

Langkah 2 Pilih Simple pada Line Charts dan kemudian Define (Gambar 3.28).



Gambar 3.28 Memilih Simple dari kotak dialog Line chart

Langkah 3 Pindahkan variabel, lalu klik Titles dan ketik judul (Gambar 3.29).



Gambar 3.29 **Define Simple Line: Summaries for group of cases** dengan variabel yang akan dipilih

Langkah 4 Klik Lanjutkan dan kemudian OK untuk mendapatkan output. Output muncul dengan judul dan garis yang mewakili tekanan darah (mmHg) dari individu dalam sumbu y dan usia orang dalam sumbu x.

Output

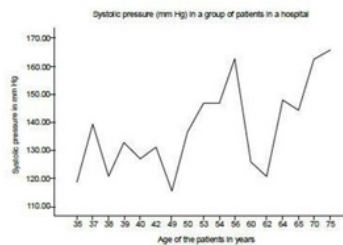


Diagram Garis Sederhana

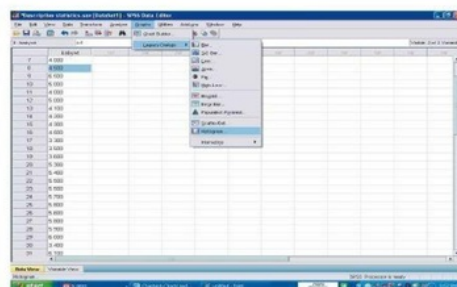
HISTOGRAM

Histogram adalah diagram dua-dimensi untuk mengukur variabel kontinyu. Ada sejumlah variabel yang ditandai pada data mahasiswa, berat dan tinggi binatang, panjang daun, dan lain lain. Dalam hal ini variabel diambil dalam sumbu X dan frekuensi variabel dalam sumbu Y. Sebuah bar atau persegi panjang dibangun pada interval kelas dan ketinggian bar sama dengan frekuensi interval kelas dan lebar batang sebanding dengan kelas interval.

Contoh 3.9

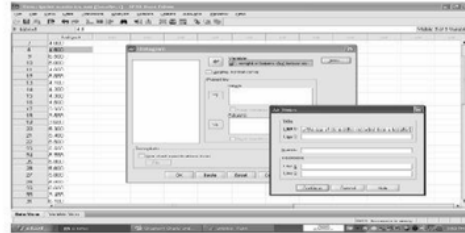
Berat badan 40 bayi yang direkam dari rumah sakit dimasukkan dalam Data Editor seperti yang dijelaskan di Bab 2. Histogram dibuat dengan langkah-langkah berikut:

Langkah 1 Pilih Grafik dari menu utama, pilih Legacy Dialog dan klik Histogram untuk membuka kotak dialog Histogram (Gambar 3.30).



Gambar 3.30 Memilih Histogram dari menu utama

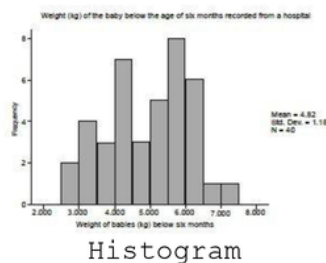
Langkah 2 Pindahkan berat bayi ke Variabel, klik Title dan ketikkan judul (Gambar 3.31).



Gambar 3.31 Histogram dengan variabel yang dipindahkan dan Kotak dialog Titles dengan judul

Langkah 3 Klik Lanjutkan, lalu klik OK untuk mendapatkan output. Output muncul dengan judul dan batang empat persegi panjang yang dibangun pada interval kelas dalam X-sumbu mulai dari 2,00 hingga 8,00, mewakili berat (kg) bayi. Ketinggian batang sama dengan frekuensi dalam sumbu Y.

Output



REVIEW LATIHAN

1. Tingkat kalsium karbonat (mg/L) dari empat kolam berbeda diberikan di bawah ini; gambarlah diagram batang dengan error bar chart.

Pond 1	Pond 2	Pond 3	Pond 4
95.5	130	124	165
100	135	125	180
102	123	130	175
112	124	124	180
115	130	135	176
112	135	124	156
113	145	142	175

2. Gambarlah box plot untuk data yang diberikan dalam latihan 1.
3. Gambarkan histogram dengan kurva frekuensi ketinggian dalam cm untuk data berikut:

155	171	170	169	167	180	158
153	172	172	175	152	163	164
177	164	165	174	164	160	165
174	169	168	165	151	175	150
164	165	163	170	167	159	
163	162	168	154	170	157	

4. Gambarlah diagram pencar dan grafik garis untuk data tinggi (cm) dan berat (kg) 10 individu, sebagai berikut:

Height (cm)	165	160	157	158	168	170	171	169	165	163
Weight (kg)	55	57	54	54	60	65	76	66	59	52

5. Pengeluaran keluarga untuk barang-barang yang berbeda diberikan di bawah ini. Sajikan data dalam bentuk diagram lingkaran.

Item	Expenditure in Rs.
Rent	2000
Food	4000
Clothing	1000
Education	900
Saving	1000
Others	1600

6. Data berikut mewakili pengeluaran tiga keluarga untuk item yang berbeda, gambarlah diagram batang yang terbagi.

Item	Expenditure in Rs		
	Family A	Family B	Family C
Rent	2000	1000	3000
Food	4000	2000	4000
Clothing	1000	1000	2000
Education	900	600	2500
Saving	1000	100	500
Others	1600	600	1600

7. Status perkawinan pria di dua tempat berbeda di kota tertentu diberikan di bawah ini.

Mewakili data dengan bagan batang (M - Married, S-Single).

Tempat 1 M, M, M, S, M, M, S, M, S, M, M, M, M, M, S, M, M, M, S, K M, M, KK MM, M, KM, M, M, M, M, M, S, M, M, M, M, M, M, M, M, M, M, S, S, S, M, M, M, M, M, M, M, M, M, M, M, M, M, M,

Tempat 2 M, M, M, M, M, M, M, M, S, M, M, S, M, S, M, S, M, M, M, KKK M, M, M, KMMMMM, M, KM S, M, M, M, M, M, M, M, M, M, M, M, M, S, S, M, M, M, M, M,

Bab 4

MEMBANDINGKAN RATA-RATA

UJI PARAMETRIK DAN NON-PARAMETRIK UNTUK MEMBANDINGKAN RATA-RATA

Bab ini membahas prosedur statistik untuk menguji signifikansi rata-rata. Tujuan dari setiap penyelidikan statistik adalah untuk menyimpulkan karakteristik suatu populasi dengan menganalisis karakteristik suatu kelompok sampel kecil. Dalam melakukan analisis ini, kita menemukan situasi yang berbeda. Dalam beberapa penelitian, kita mengasumsikan bahwa sampel/ beberapa sampel berasal dari populasi yang terdistribusi normal dan memiliki rata-rata hitung dan varians yang sama, misalnya, tinggi individu dalam suatu populasi. Untuk menguji signifikansi perbedaan, hipotesis diajukan pada parameter populasi. Prosedur pengujian semacam itu disebut uji parametrik. Banyak prosedur statistik umum lainnya seperti regresi dan analisis varian memiliki asumsi yang sama.

Namun, ada metode/uji statistik yang terdiri dari prosedur yang tidak memerlukan estimasi rata-rata populasi dan/atau tidak menyatakan Hipotesis tentang parameter, misalnya, peningkatan denyut nadi dalam melakukan beberapa kegiatan petualangan. Dalam hal ini kita tidak memiliki populasi yang ada sehingga kita tidak mengasumsikan mean dan varians (parameter populasi) dalam mengajukan hipotesis.

Prosedur yang tidak menyatakan hipotesis tentang parameter disebut uji non-parametrik. Penjelasan dan contoh berikut berhubungan dengan uji-t "Student" dan non-parametrik yang setara seperti Uji Mann-Whitney dan Uji Wilcoxon.

UJI-T (STUDENT t-test)

A.S. Gosset menggambarkan distribusi yang disebut distribusi-t dan uji signifikansi berdasarkan itu disebut uji-t. Distribusi t didasarkan pada derajat kebebasan. Derajat kebebasan didefinisikan sebagai jumlah varian yang dapat dimasukkan dalam distribusi tersebut sebelum nilai dari sisa varian ditetapkan untuk menghasilkan nilai tertentu ($df = 11 - 1$). Uji-t memungkinkan kita untuk menguji signifikansi perbedaan antara dua rata-rata sampel atau signifikansi dari rata-rata tunggal. Prosedur ini masing-masing disebut uji dua sampel dan uji satu sampel.

UJI DUA SAMPEL

Uji-t sampel independen.

Pengambilan keputusan yang aman tergantung pada pemahaman data yang cermat. Ada berbagai penelitian eksperimental dan quasi-eksperimental di mana pengamatan variabel yang sama dibandingkan dalam kondisi yang berbeda. Dalam situasi seperti itu uji dua sampel adalah uji yang paling tepat. Ketika kita melanjutkan dengan uji dua sampel, kita harus memutuskan apakah kedua kelompok itu independen atau berhubungan. Apa yang kita maksudkan dengan sampel independen? Sampel 100 individu diambil dari populasi dan dibagi secara acak menjadi dua kelompok dan satu kelompok dikenakan beberapa kondisi eksperimental dan sisanya untuk mengontrol kondisi. Dalam pengaturan ini, individu dalam satu kelompok tidak berpengaruh pada kelompok lain. Kedua kelompok bersifat independen, yaitu, kedua kelompok adalah sampel data yang independen, masing-masing terdiri dari 50 pengamatan. Uji yang paling tepat untuk situasi ini adalah uji-t sampel independen.

Prosedur uji-t sampel independen adalah cara membandingkan dua kelompok kasus. Idealnya, untuk uji ini, subyek harus secara acak ditugaskan ke dua kelompok, sehingga perbedaan dalam respons adalah karena perlakuan (atau kurangnya perlakuan) dan bukan karena faktor lain. Ini tidak terjadi jika Anda membandingkan pendapatan rata-rata untuk pria dan wanita. Seseorang tidak secara acak ditugaskan untuk menjadi pria atau wanita. Jika Anda ingin menerapkan uji-t sampel independen, Anda harus memastikan bahwa perbedaan dalam faktor-faktor lain tidak menutupi atau meningkatkan perbedaan yang berarti. Perbedaan dalam pendapatan rata-rata dapat dipengaruhi oleh faktor-faktor seperti pendidikan (dan bukan oleh jenis kelamin saja).

Sebagai contoh, pasien dengan tekanan darah tinggi secara acak dimasukkan ke kelompok plasebo dan kelompok perlakuan. Subyek plasebo menerima pil tidak aktif dan subyek pengobatan menerima obat baru yang diharapkan dapat menurunkan tekanan darah. Setelah subyek diberi pil selama dua bulan, uji-t dua sampel digunakan untuk membandingkan tekanan darah rata-rata untuk kelompok plasebo dan kelompok perlakuan. Setiap pasien termasuk dalam satu kelompok dan diuji satu kali.

Ada banyak situasi di mana peneliti harus memutuskan apakah perbedaan yang diamati antara dua rata-rata sampel dikaitkan

dengan kebetulan atau apakah kedua sampel tersebut berasal dari dua populasi dengan rata-rata yang tidak sama. Misalnya, kita mungkin ingin tahu apakah ada perbedaan nyata dalam kinerja pria dan wanita dalam menyelesaikan tugas sehubungan dengan waktu. Orang mungkin tertarik untuk mengetahui apakah diet rata-rata di satu negara lebih bergizi daripada di negara lain. Seorang ahli biologi mungkin bersedia untuk mengetahui apakah perbedaan panjang sayap dari varietas burung tertentu, pada dua wilayah geografis yang berbeda, cukup besar untuk menyarankan spesiasi. Seorang penyedia layanan kesehatan mungkin tertarik untuk merekomendasikan diet tertentu untuk meningkatkan kadar hemoglobin pada anak-anak. Dalam semua situasi ini, kita memiliki dua set data dari dua populasi yang berbeda, baik hipotesis atau yang ada. Di sini hipotesis nol adalah bahwa tidak ada perbedaan dalam rata-rata dua sampel.

Rumus untuk perhitungan manual adalah sebagai berikut:

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{SE_1^2 + SE_2^2}}$$

dimana,

\bar{X}_1 — rata-rata sample 1

\bar{X}_2 — rata-rata sample 2

SE_1 — standard error sample 1

SE_2 — standard error sample 2

Standar Error dihitung dengan menerapkan rumus berikut, di mana SD adalah standar deviasi dan n_1 dan n_2 adalah ukuran sampel 1 dan 2.

$$SE_1 = \frac{SD_1}{n_1}; \quad SE_2 = \frac{SD_2}{n_2}$$

Simpangan baku dihitung sebagai

$$SD_1 = \sqrt{\frac{\sum (X_i - \bar{X}_1)^2}{n_1 - 1}}; \quad SD_2 = \sqrt{\frac{\sum (X_i - \bar{X}_2)^2}{n_2 - 1}}$$

dimana,

45

n1 – ukuran sampel 1

n2 – ukuran sampel 2

Value t yang dihitung mengikuti $n1 + n2 - 2$ derajat kebebasan. Hipotesis nolnya adalah. $\bar{X}_1 = \bar{X}_2$

Value-t yang dihitung mungkin positif atau negatif.

Tingkat signifikansi ditetapkan sebagai 0,05 atau 0,01. Terserah Anda untuk mengatur salah satu level ini. Dalam kasus sebelumnya kita menolak hipotesis nol jika ini lebih besar dari persentil ke-97,5 atau kurang dari persentil ke-2,5 dari distribusi-t dengan derajat kebebasan tertentu. Ketika kita mengatakan tingkat signifikansi adalah 0,05, tingkat kepercayaan 95% dan jika tingkat signifikansi 0,01, tingkat kepercayaan 99%.

Uji-t sampel independen dengan SPSS

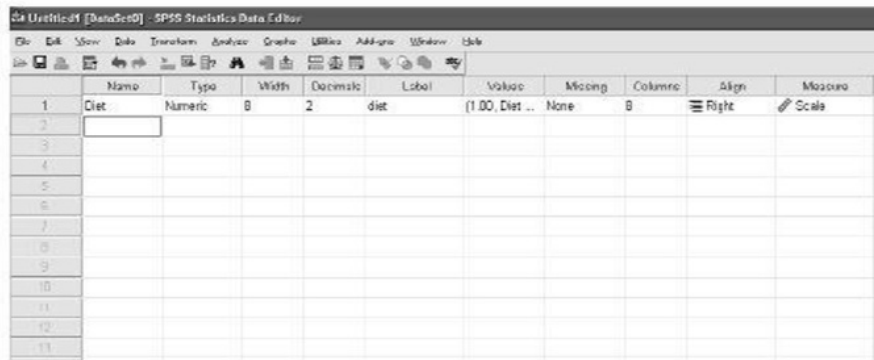
Contoh 4.1

Dua kelompok anak diberi makan pada dua diet yang berbeda yaitu, diet A dan diet B. Tingkat hemoglobin dalam darah diperkirakan dan disajikan pada tabel di bawah ini. Ujilah keunggulan diet B dibandingkan diet A dalam meningkatkan kadar hemoglobin.

Diet	Haemoglobin (gm%)													
A	10	11	12	11	10	9	10	9.5	12	13	11	11	12	10
B	11	11	12	13	11	9	12	10	11	12	12	12	13	10

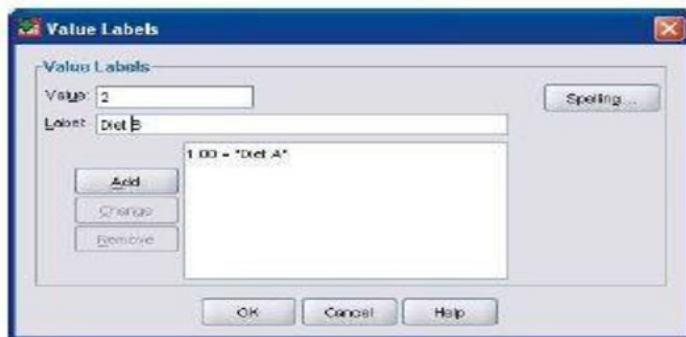
Hipotesis nol: Diet B tidak lebih baik daripada diet A dalam meningkatkan kadar hemoglobin.

Langkah 1 Buka Data Editor, klik Variabel View, beri nama variabel sebagai "Diet" pada Name, pilih Numeric pada Type, label sebagai Diet pada Label (Gambar 4.1).



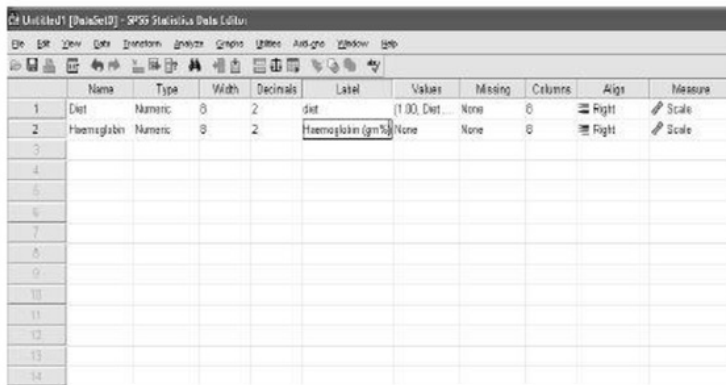
Gambar 4.1 Variabel view dengan nama variabel Diet yang diketik pada Name

Langkah 2 Klik pada area abu-abu di bawah Value. Jendela *pop up* terbuka seperti pada Gambar 4.2, ketik 1 di kotak Value dan "Diet A" di kotak Label dan klik **Add** untuk dipindah ke kotak di bawah. Ketik "2" di kotak Value dan "Diet B" di kotak Label dan klik Add untuk dipindah ke kotak di bawah dan klik OK.



Gambar 4.2 Kotak Value Label untuk memasukkan nomor kode untuk kelompok variabel

Langkah 3 Lanjut ke baris kedua dan ketik "Hemoglobin" pada Name pilih **N**umerik pada Type, ketik "Level hemoglobin (gm%)" pada Label (Gambar 4.3).



Gambar 4.3 Variabel View dengan dua variabel dinamai pada Name

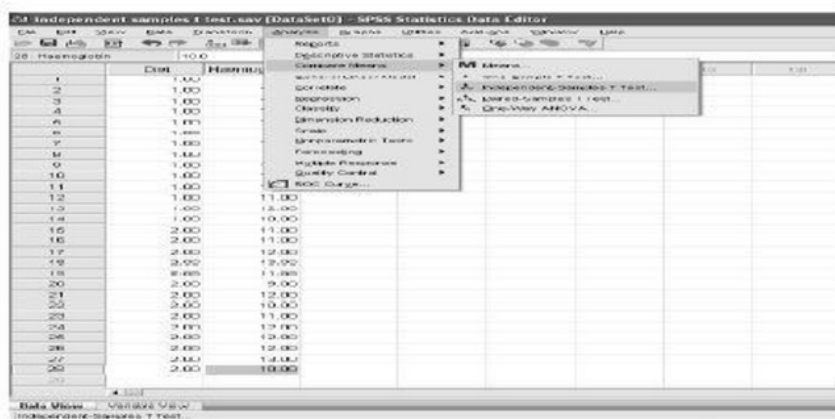
Langkah 4 Klik Data View dan masukkan data dalam kolom pada Diet sebagai "1" untuk 1 hingga 14 kasus dan "2" dari 15 hingga 28 kasus (di sini diet dimasukkan sebagai variabel kategori di mana 1 mewakili diet A dan 2 mewakili diet B). Pada kolom ke dua, masukkan nilai hemoglobin di bawah Hemoglobin seperti pada Gambar 4.4.

The screenshot shows the SPSS Data View for the same dataset. It displays data for 28 cases across two variables: Diet and Hemoglobin.

	Diet	Hemoglobin
1	1.00	10.00
2	1.00	11.00
3	1.00	12.00
4	1.00	11.00
5	1.00	10.00
6	1.00	9.00
7	1.00	10.00
8	1.00	9.50
9	1.00	12.00
10	1.00	13.00
11	1.00	11.00
12	1.00	11.00
13	1.00	12.00
14	1.00	10.00
15	2.00	11.00
16	2.00	11.00
17	2.00	12.00
18	2.00	13.00
19	2.00	11.00
20	2.00	9.00
21	2.00	12.00
22	2.00	10.00
23	2.00	11.00
24	2.00	12.00
25	2.00	12.00
26	2.00	13.00
27	2.00	13.00
28	2.00	10.00

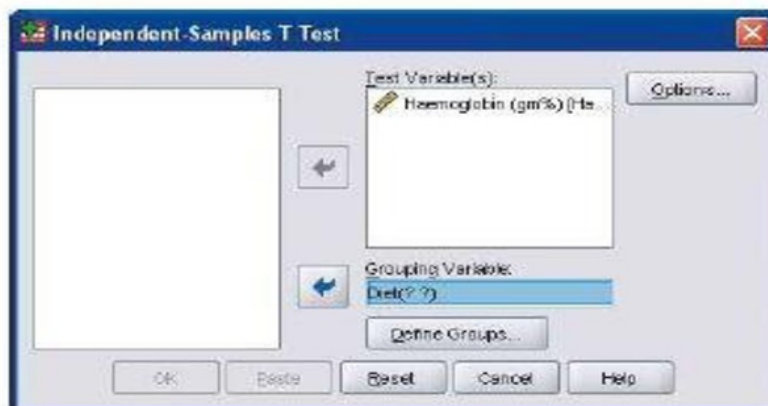
Gambar 4.4 Data View dengan data yang dimasukkan untuk dua variabel

Langkah 5 Pilih **Analyze**, pilih **Compare Means** dan kemudian pilih **Independent Sampel T-test** (Gambar 4.5).



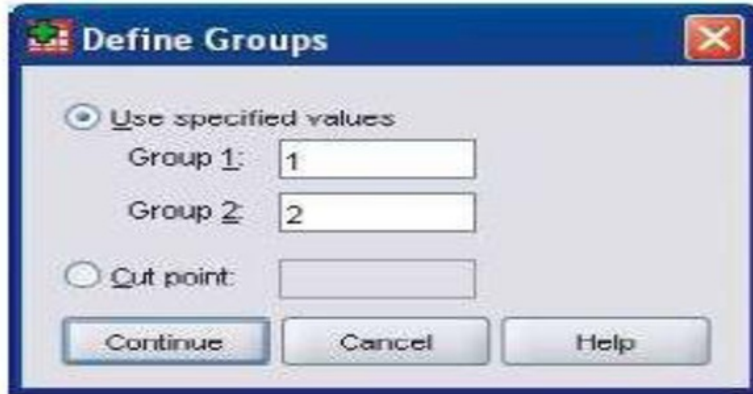
Gambar 4.5 Memilih Independen-sampel T-test dari menu

Langkah 6 Jendela Independent Samples T-test terbuka, pindahkan Variabel Haemoglobin ke Test Variable(s) dan Variabel Diet ke Grouping Variable: (Gambar 4.6).



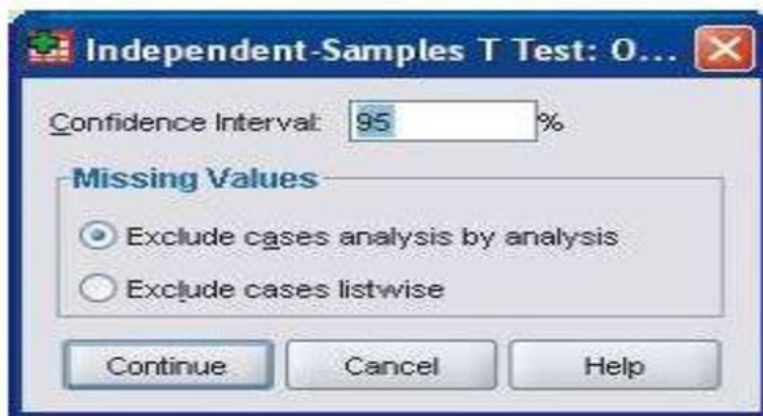
Gambar 4.6 Independent Sample T-test dengan variabel yang pindahkan ke Test Variabel dan ke kotak Grouping Variabel

Langkah 7 Klik pada Define Groups, jendela pop up dengan Define Groups terbuka. Ketik 1 di Grup 1 dan "2" di Grup 2 (di sini kita memberikan nilai 1 dan 2 karena kita telah menentukan nilai-nilai ini untuk dua kelompok yang akan dibandingkan, masing-masing Diet A dan Diet B), klik Lanjutkan (Gambar 4.7) .



Gambar 4.7 Define Groups dengan nilai yang ditentukan

Langkah 8 Klik Opsi. Jendela pop up terbuka (Gambar 4.8). Centang kotak Interval Batas kepercayaan. Ini memiliki nilai default 95%. Klik Lanjutkan (Jika Anda ingin meningkatkan batas kepercayaan hingga 99%, ketikkan di sana sebagai "99").



Gambar 4.8 Independent samples T-test: dengan batas interval kepercayaan 95% dipilih

Langkah 9 Klik OK untuk menjalankan analisis. Output muncul seperti di bawah ini (output 1 dan 2).

Output 1

Group Statistics

	diet	N	Mean	Std. Deviation	Std. Error Mean
Haemoglobin (gm%)	Diet A	14	10.8214	1.14	0.3
	Diet B	14	11.3571	1.15	0.31

Output 2

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference		Lower	Upper
Haemoglobin (gm%)	Equal variances assumed	.002	.966	-1.239	26	.226	-0.54	0.43		-1.42	0.35
	Equal variances not assumed			-1.239	25.906	.226	-0.54	0.43		-1.42	0.35

Interpretasi

Output 1 memberikan tingkat hemoglobin rata-rata pada individu yang diberi Diet A adalah 10.82 dengan standar deviasi 1,136 (gm%) dan untuk Diet B, tingkat rata-rata adalah 11,357 dengan standar deviasi 1,15 (gm%).

Output 2 memberikan nilai-t, derajat kebebasan, tingkat signifikansi dan batas interval kepercayaan 95% untuk rata-rata. Nilai-t -1.239 untuk 26 df (14 + 14 - 2 karena setiap kelompok memiliki nilai 14) tidak signifikan karena nilai signifikansi (untuk Uji dua sisi) adalah 0,226 yang > 0,05. Oleh karena itu, kita TIDAK MENOLAK Hipotesis nol, yaitu, diet B tidak lebih baik daripada diet A dalam meningkatkan kadar hemoglobin.

Uji t-Berpasangan (Paired sample t-test)

Dalam Uji - t sampel independen kita telah menguji signifikansi perbedaan antara rata-rata dua sampel independen. Dalam hal ini pengamatan atau nilai setiap item sepenuhnya independen. Tetapi dalam situasi di mana pengamatan tidak sepenuhnya independen

tetapi tidak bebas, mereka diberikan sebagai pasangan pengamatan (masing-masing pasangan dari subyek yang sama). Terkadang, kita memiliki satu set data, di mana pengamatan atau pengukuran dilakukan pada orang yang sama, misalnya, seorang dokter ingin menguji efisiensi obat tertentu untuk mengurangi tekanan darah. Dalam hal ini ia dapat mengambil kelompok 50 orang yang hipertensi dan mengukur tekanan darah dan memilikinya sebagai satu set data. Ke-50 orang ini diberikan obat tertentu dan tekanan darah diukur dan nilainya ditulis terhadap setiap kasus. Sekarang kita memiliki satu set pasangan data, mis., tekanan darah sebelum dan sesudah memberikan obat. Dalam kasus ini juga kita memiliki dua set pengamatan tetapi sepasang untuk setiap kasus (Meskipun ada 100 pengamatan dalam dua kelompok, mereka didasarkan hanya pada 50 individu). Uji yang paling tepat adalah uji dua sampel, tetapi di sini merupakan Uji-t sampel berpasangan dan bukan Uji-t sampel independen.

Prosedur Uji-berpasangan-sampel membandingkan rata-rata dua variabel untuk satu kelompok. Prosedur menghitung perbedaan antara nilai-nilai dari dua variabel untuk setiap kasus dan menguji apakah rata-rata berbeda dari 0.

Formula untuk perhitungan manual adalah:

$$t = \frac{\bar{D}}{S / \sqrt{n}}$$

dimana,

\bar{D} —rata-rata perbedaan

S —standar deviasi

n — Jumlah sampel

$$\bar{D} = \frac{\sum D}{n}$$

$$D = X_i - Y_i$$

dimana,

D — perbedaan antara pasangan pengamatan

X_i — Value sebelum

Y_i — Value sesudah

$$S = \sqrt{\frac{(\sum D - \bar{D})^2}{n}}$$

Paired sample t-Uji-t dengan SPSS

Contoh 4.2

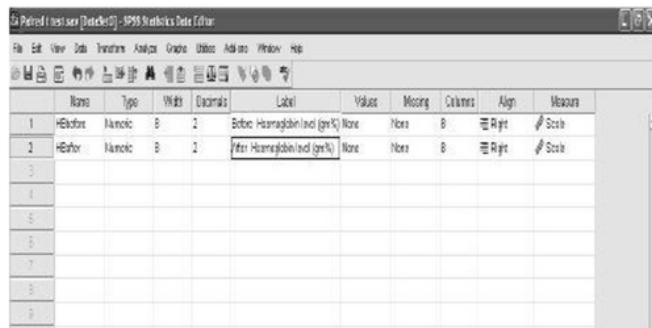
Seorang peneliti ingin mengevaluasi efek dari diet tambahan tertentu dalam meningkatkan tingkat hemoglobin pada manusia. Peneliti memilih sekelompok 15 orang, tingkat hemoglobin pada orang-orang ini diperoleh dan kemudian orang-orang ini diberi makanan tambahan. Setelah makan untuk jangka waktu yang cukup, tingkat hemoglobin pada orang-orang ini dihitung. Data yang diperoleh dalam penelitian ini diberikan dalam bentuk tabel. Hitunglah apakah diet tambahan efektif dalam meningkatkan kadar hemoglobin (gm%)?

Before	10	12	11	10	9.5	10.5	11.5	10	9.8	10.5	12	11	11.5	10.5	9.9
After	11	12	12	11	10	11.5	12	10	10	11	13	12	12	11	10

Hipotesis nol: Diet tambahan tidak efektif dalam meningkatkan kadar hemoglobin pada manusia.

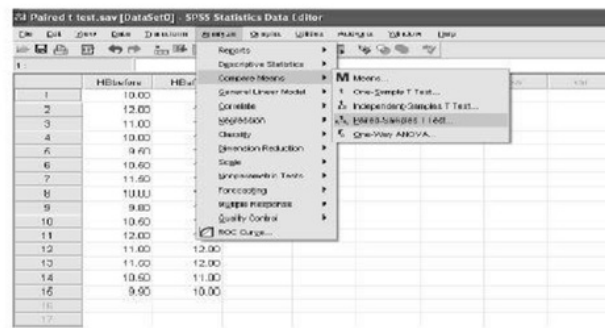
Langkah 1 Buka Data Editor, klik Variabel View, beri nama variabel "HBbefore" di bawah Name, pilih Numerik pada Type, beri label "Before-hemoglobin level (gm%)" pada Label, lalu ke baris kedua dan ketik di bawah Name sebagai "HBafter", pilih Numeric pada Type, beri label "After-Hemoglobin level (gm%)" pada Label (Gambar 4.9).

Langkah 2 Klik Data view dan masukkan data di bawah Sebelum dan Sesudah.



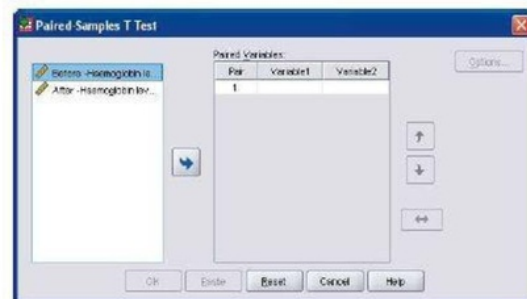
Gambar 4.9 Variabel view dengan dua variabel yang dimasukkan

Langkah 3 Pilih Analyze, lalu pilih Compare Means, lalu pilih Paired-Samples T-test (Gambar 4.10).



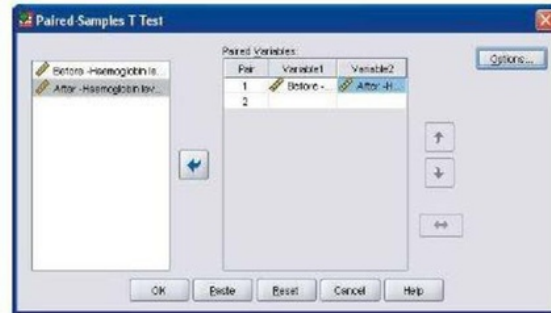
Gambar 4.10 Pilihan Paired-Samples T-test dari menu utama

Langkah 4 Jendela pop-up muncul sebagai Paired-Sample T-test seperti pada Gambar 4.11.



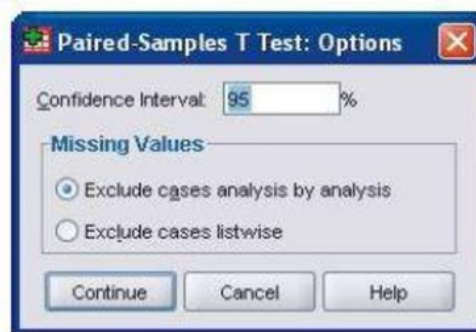
Gambar 4.11 Paired-Sample T-test

Langkah 5 Pindahkan variabel Before Hemoglobin ke Variabel 1 dan After - Hemoglobin ke Variabel 2 seperti pada Gambar 4.12.



Gambar 4.12 Paired-Sample T-test dengan variabel berpasangan

Langkah 6 Klik Opsi dan centang Confidence Interval (Gambar 4.13). Nilai default sebagai 95% (Jika Anda ingin meningkatkan tingkat kepercayaan ke 99 ketik "99"), klik Lanjutkan.



Gambar 4.13 Paired-Samples T-test: Pilihan dengan Confidence Interval 95%

Langkah 7 Klik OK untuk menjalankan analisis. Output tampak seperti di bawah ini:

Output 1

Paired Samples Statistics				
	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 Before -Haemoglobin level (gm%)	10.6467	15	.81404	.21019
After -Haemoglobin level (gm%)	11.2333	15	.94239	.24332

Output 2

Paired Samples Test									
		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	Before-Haemoglobin level (gm%) - After-Haemoglobin level (gm%)	-58657	39073	.10088	-80304	-37029	-5.815	.000	

Interpretasi

Output 1 tingkat hemoglobin rata-rata pada individu sebelum memakan diet tambahan adalah 10,64 dengan standar deviasi 0,81 (gm%) dan 11,23 dengan standar deviasi 0,94 (gm%) setelah memakan diet tambahan.

Output 2 memberikan nilai $-t$, derajat kebebasan, tingkat signifikansi dan interval kepercayaan 95% untuk perbedaan rata-rata. Nilai t -5,815 untuk 14 derajat kebebasan (df) sangat signifikan karena value signifikan untuk uji dua-ekor adalah 0,000 (Output 2). Oleh karena itu, kita MENOLAK Hipotesis nol. Dengan demikian, diet tambahan efektif dalam meningkatkan kadar hemoglobin pada manusia.

58

Uji-t Satu Sampel

Prosedur Uji-t satu sampel menguji apakah rata-rata hitung dari variabel tunggal berbeda dari sebuah konstanta tertentu. Uji ini mengasumsikan bahwa data terdistribusi secara normal. Prosedur ini menguji perbedaan antara rata-rata hitung sampel dan nilai yang diketahui atau di hipotesiskan. Misalkan kita ingin menguji:

1. jika rata-rata hitung sampel berbeda secara signifikan dari nilai rata-rata populasi yang ditentukan
2. jika populasi yang diberikan memiliki nilai tertentu dari populasi rata-rata hitung, misalnya A_0
3. jika sampel acak x_1, x_2, \dots, x_n dari sejumlah n diambil dari populasi normal dengan rata-rata hitung A_0 .

Pada dasarnya, ketiga situasi ini membutuhkan prosedur yang sama. Dalam kondisi ini Hipotesis nol ditetapkan sebagai berikut:

i. H_0 : Tidak ada perbedaan yang signifikan antara rata-rata hitung sampel dan rata-rata hitung populasi, yaitu, dengan kata lain $\bar{X} = \mu_0$

ii. H_0 : yaitu, $U = U_0$ rata-rata hitung populasi adalah U_0 .

iii. H_0 : Sampel acak yang diberikan diambil dari populasi normal dengan rata-rata hitung.

$$t = \frac{|\bar{X} - \mu_0|}{S / \sqrt{n}}$$

dimana,

S – standar deviasi sampel

n – ukuran sampel

$$S = \sqrt{\frac{1}{n-1} \sum (X_i - \bar{X}_i)^2}$$

dimana,

X_i – Value dari variabel X, dan

\bar{X} – rata-rata hitung

Aplikasi dari Uji-t satu sampel

Beberapa variabel seperti tinggi badan, berat badan, panjang daun, output pernafasan, denyut nadi, tekanan darah, gula darah, kadar hemoglobin, berat buah anggur, buah-buahan di kebun, dan lain-lain, adalah variabel yang dapat diklasifikasikan dan diberikan pada skala interval. Ketika nilai-nilai ini diplot pada grafik, grafik yang dihasilkan paling sering berbentuk lonceng (bell-shaped) dengan modus tunggal dan disebut sebagai kurva normal. Nilai-nilai ukuran yang berbeda dari tendensi sentral seperti rata-rata hitung, median dan modus cenderung terjadi di pusat distribusi. Nilai-nilai variabel di kedua sisi ukuran tendensi sentral, yaitu rata-rata didistribusikan secara merata. Jumlah penyimpangan nilai yang berbeda dari rata-rata di kedua sisi sama dengan nol. Distribusi semacam ini disebut

distribusi normal. Uji-t Satu-sampel berlaku untuk setiap data pada skala kontinyu dan peneliti mengumpulkan data hanya pada satu sampel dan ukuran sampel besar. Uji ini tidak berlaku untuk sampel kecil. Uji-t Satu sampel memungkinkan kita untuk menguji apakah rata-rata hitung sampel (dari variabel interval terdistribusi normal) secara signifikan berbeda dari nilai yang Hipotesis.

Dalam semua contoh di atas, berdasarkan data sampel, Value tunggal ditetapkan ke parameter populasi. Parameter populasi, sehingga ditugaskan diterima atau ditolak oleh pengujian Hipotesis dan kesimpulan ditarik pada rata-rata hitung populasi dan standar deviasi. Penerapan Uji t satu sampel dijelaskan dengan contoh-contoh berikut.

Uji-t Satu-sampel dengan SPSS

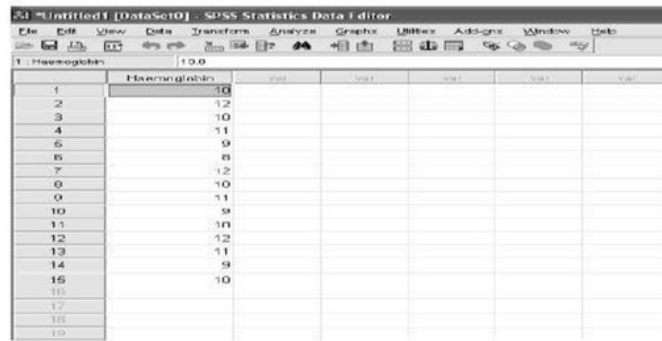
Contoh 4.3

Data berikut memberikan kadar hemoglobin (gm%) dalam kelompok 15 mahasiswi yang belajar di perguruan tinggi. Ujilah apakah kadar hemoglobin pada wanita muda secara signifikan lebih rendah dari tingkat rata-rata 13% gm yang diharapkan untuk populasi wanita. Selidiki apakah wanita dalam kelompok usia memiliki kecenderungan anemia.

Case	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Haemoglobin (gm%)	10	12	11	10	9	8	12	10	11	9	10	12	11	9	10

Hipotesis nol: Kadar hemoglobin pada wanita yang belajar di perguruan tinggi tidak berbeda secara signifikan dari tingkat hemoglobin 13 gm% dari populasi wanita.

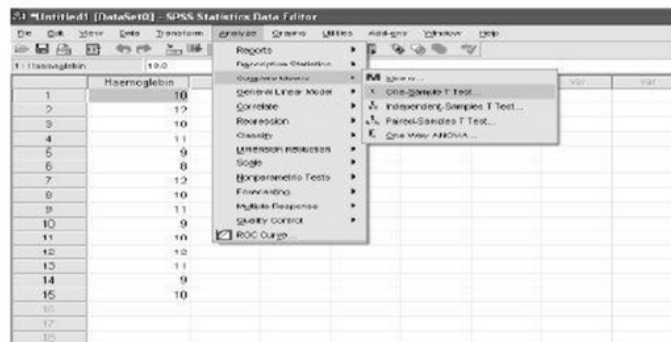
Langkah 1 Buka Data Editor, klik Variabel view, beri nama variabel sebagai "Hemoglobin", pilih Numerik pada Type, beri label sebagai "Hemoglobin (gm%)" pada Label, klik Data view dan masukkan data dalam Data view dari data Hemoglobin seperti pada Gambar 4.14.



	Hemoglobin	Var1	Var2	Var3	Var4
1	10				
2	12				
3	10				
4	11				
5	9				
6	8				
7	12				
8	10				
9	11				
10	9				
11	10				
12	12				
13	11				
14	9				
15	10				
16					
17					
18					

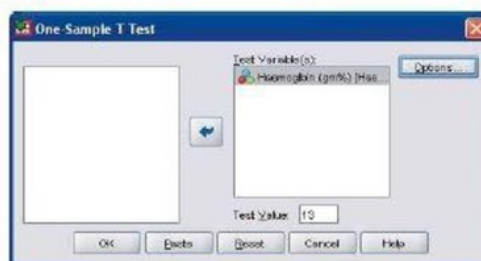
Gambar 4.14 Data view dengan nilai hemoglobin

Langkah 2 Pilih **Analyze**, pilih **Compare Means**, lalu pilih **One-sample T-test** (Gambar 4.15).



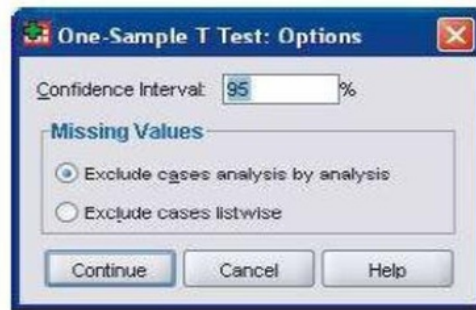
Gambar 4.15 Seleksi One-Sample T-test dari menu utama

Langkah 3 Pindahkan variabel yang diteliti "Hemoglobin (gm%)" ke **Test Variable(s)**: (Gambar 4.16). Ketik pada kotak **Test Value** :13 (Di sini kita menguji apakah kadar rata-rata hemoglobin sampel ini berbeda secara signifikan dari kadar normal 13% gm yang diharapkan untuk wanita dalam populasi).



Gambar 4.16 Uji-T satu sampel dengan variabel yang dipindah dalam kotak uji variabel

Langkah 4 Klik Opsi dan centang kotak Confidence Interval (Gambar 4.17). Nilai default: 95%. Klik Lanjutkan (Jika Anda ingin meningkatkan batas kepercayaan menjadi 99% ketik "99").



Gambar 4.17 Uji-t Satu Sampel: Pilihan Interval kepercayaan

Langkah 5 Klik OK untuk menjalankan analisis. Output tampak seperti yang di bawah ini.

Output 1

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
Haemoglobin (gm%)	15	10.27	1.22	0.32

Output 2

One-Sample Test						
	Test Value = 13					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Haemoglobin (gm%)	-8.66	14	0.000	-2.73	-3.41	-2.06

Interpretasi

Output 1 kadar hemoglobin rata-rata 10,27 (gm%) dan standar deviasi 1,22.

Output 2 Nilai t, derajat kebebasan (df),signifikansi dan interval kepercayaan 95% rata-rata hitung nilai-t = -8,657 untuk df = 14 sangat signifikan karena nilai signifikan adalah 0,000. Oleh karena itu, kita MENOLAK Hipotesis nol. Dengan demikian, kadar hemoglobin pada wanita muda secara signifikan lebih rendah

dari kadar rata-rata 13gm% yang diharapkan untuk wanita dalam populasi. Para wanita dalam kelompok usia tersebut memiliki kecenderungan anemia.

Contoh 4.4

Berikut ini adalah data tentang peningkatan denyut nadi (denyut /menit) yang dicatat oleh dokter pada 32 orang saat melakukan tugas yang diberikan. Apakah ada peningkatan signifikan dalam denyut nadi saat melakukan tugas ini.

Increase in pulse rate (beats/minute)							
27	25	19	28	35	23	24	22
14	30	32	34	23	26	29	27
27	24	31	22	23	38	25	16
32	29	26	25	28	26	21	28

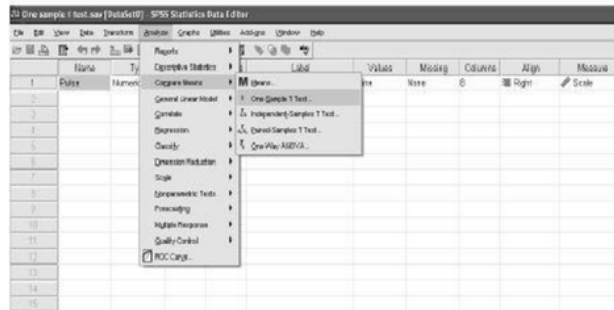
Hipotesis Nol: Tidak ada peningkatan denyut nadi.

Langkah 1 Buka Data Editor, Klik Variabel view, beri nama variabel sebagai Nadi, pilih Numeric pada Type, beri label "Pulse rate (beats/minute)" pada Label. 20 klik Data View dan masukkan data dalam Data View pada Nadi seperti yang terlihat pada Gambar 4.18.

Case	Pulse rate (beats/minute)
1	27
2	25
3	19
4	28
5	35
6	23
7	24
8	22
9	14
10	30
11	32
12	34
13	23
14	26
15	29
16	27
17	27
18	24
19	31
20	22
21	23
22	38
23	25
24	16
25	32
26	29
27	26
28	25
29	28
30	26
31	21
32	28

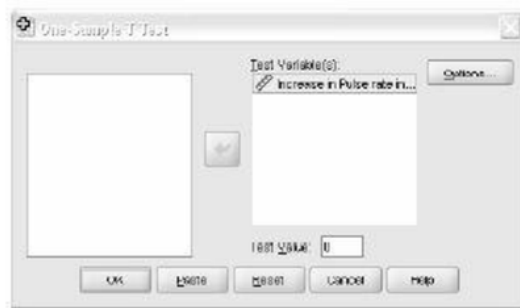
Gambar 4.18 Data Editor dengan nilai yang dimasukkan

Langkah 2 Pilih Analyze, Compare Means, lalu pilih One-Sample T-test (Gambar 4.19).



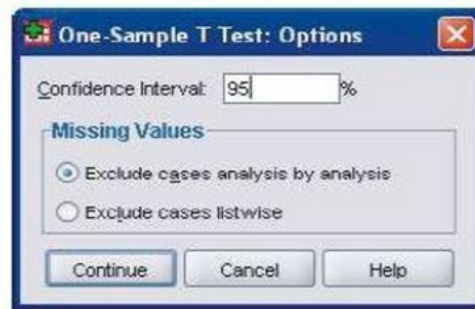
Gambar 4.19 Pilihan Uji-t Satu Sampel dari menu utama

Langkah 3 Pindahkan variabel yang diteliti (Hemoglobin) ke kotak **test variable(s)**: (Gambar 4.20). Dalam kotak **Test Value** Anda akan melihat 0 sebagai default, jangan mengubahnya (Di sini kita tidak memiliki peningkatan denyut nadi yang diharapkan. Hal ini berbeda dengan situasi pada contoh sebelumnya dari kadar hemoglobin, di mana kita memiliki kadar yang diharapkan untuk wanita dalam populasi).



Gambar 4.20 Uji-t Satu Sampel dengan variabel yang dipindah dalam kotak Test Variable

Langkah 4 Klik Opsi **One-Sample T-test: Opsi, Confidence Interval 95%** sebagai default (jika tidak ketik 95 di dalamnya) dan kemudian klik Lanjutkan (Gambar 4.21).



Gambar 4.21 Pilihan Interval kepercayaan Uji-t Satu Sampel

Langkah 5 Klik OK untuk menjalankan analisis. Output muncul seperti pada Output 1 dan 2.

Output 1

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
Increase in Pulse rate in beats / minutes	32	26.5313	5.43	0.96

Output 2

One-Sample Test						
	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
Increase in Pulse rate in beats / minutes	27.64	31	0.000	26.53125	Lower 24.5736	Upper 28.4889

Interpretasi

Output 1 memberikan peningkatan rata-rata dalam denyut nadi 26,5 denyut/menit dan standar deviasi 5,43.

Output 2 memberikan nilai t, derajat kebebasan, tingkat signifikansi dan interval kepercayaan 95% untuk rata-rata hitung. Nilai t = 27,64 untuk df = 31 sangat signifikan karena nilai signifikansi untuk uji dua-ekor adalah 0,000. Oleh karena itu, ada peningkatan denyut nadi orang yang diukur tersebut. Peningkatan rata-rata dalam denyut nadi (yaitu rata-rata populasi) terletak pada interval dari 24,57 hingga 28,489 denyut /menit. Untuk membuatnya sederhana, peningkatan rata-rata dalam

denyut nadi populasi mungkin terletak antara 24,57 dan 28,049 denyut/menit dengan keyakinan 95%.

UJI LAIN UNTUK MEMBANDINGKAN NILAI RATA-RATA

Uji non-parametrik seperti prosedur pengujian dua sampel kadang-kadang dinyatakan berlaku untuk data pada skala ordinal, tetapi ini tidak demikian. Ini berlaku untuk data pada skala interval atau rasio.

Uji Mann-Whitney untuk Sampel Independen

Uji Mann-Whitney adalah analog non-parametrik untuk Uji sampel independen. Ini adalah salah satu uji signifikansi non-parametrik yang paling dikenal. Awalnya diusulkan oleh Frank Wilcoxon pada tahun 1945, untuk ukuran sampel yang sama, dan diperluas ke ukuran sampel yang berubah-ubah oleh Mann dan Whitney (1947). Seperti Uji non-parametrik lainnya, menggunakan data peringkat ketimbang nilai mentah untuk menghitung parameter statistik. Ketika mengerjakan masalah secara manual, dengan pemeringkatan untuk seluruh data, peringkat Mann-Whitney (U) dihitung dengan menerapkan rumus

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

di mana, n_1 dan n_2 adalah jumlah observasi pada sampel 1 dan 2, R_1 adalah jumlah dari jajaran observasi dalam sampel 1 dan 2.

Dalam Uji ini U dihitung ditafsirkan berdasarkan nilai-nilai tabel seperti pada Uji lainnya. Karena menggunakan SPSS, detailnya tidak dilengkapi di sini. Uji Mann-Whitney adalah salah satu dari Uji non-parametrik yang paling kuat.

Uji Mann-Whitney U membutuhkan empat kondisi:

1. Variabel dependen harus pada skala interval atau rasio atau setidaknya pada skala ordinal.
2. Variabel independen hanya memiliki dua level.
3. Desain studi adalah antara-subyek.
4. Subyek tidak dicocokkan lintas kondisi.

Penerapan Uji Mann-Whitney

Uji Mann-Whitney digunakan di berbagai bidang, tetapi sering kali digunakan dalam bidang seperti Psikologi², Kedokteran, Keperawatan, dan Bisnis. Sebagai contoh, dalam Psikologi, Uji Mann-Whitney digunakan untuk membandingkan sikap, perilaku, dll. Dalam kedokteran², digunakan untuk mengetahui efek dari dua obat. Ini juga digunakan untuk mengetahui apakah obat tertentu menyembuhkan suatu penyakit. Dalam Bisnis, ini dapat digunakan untuk mengetahui preferensi orang yang berbeda.

Mann-Whitney dengan SPSS Sampel independen pada skala interval.

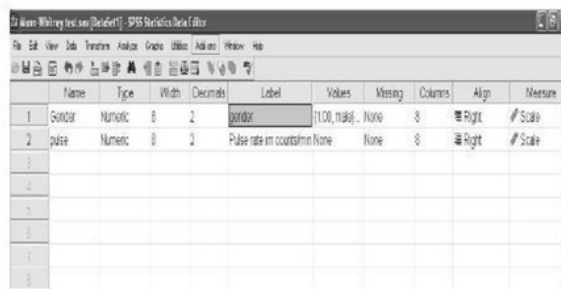
Contoh 4.5

Denyut nadi (denyut/menit) dari dua sampel acak independen, satu dari populasi pria dan yang lainnya dari populasi wanita yang duduk dalam latihan¹¹ yang bergerak cepat diberikan di bawah ini. Hitunglah apakah ada perbedaan dalam denyut nadi rata-rata dua (pria dan wanita) populasi.

Male	72	69	72	68	68	70	70	68	68	69	69	75	74
Female	73	72	75	70	68	73	74	74	70	72	71	71	63

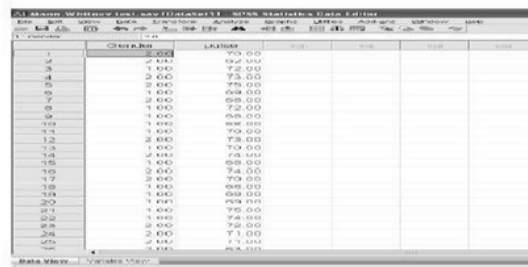
Hipotesis nol: Tingkat denyut nadi laki-laki dan perempuan yang duduk dalam latihan yang bergerak cepat adalah sama.

Langkah 1 Buka Data Editor, klik Variabel view, beri nama variabel "Gender", pilih Numerik di bawah Type, beri label sebagai "Gender" di bawah Label (Gambar 4.22). Klik pada area abu-abu pada Value dan ketik "1" di kotak Value dan "male" di kotak Label dan klik Add untuk pindahkan ke kotak di bawahnya, ketik "2" di kotak Value dan "female" di kotak Label, klik Add untuk dipindah dan klik OK. Lalu ke baris kedua dan ketik "Pulse" pada Name, pilih Numeric pada Type, ketik "Pulse rate (counts/minute)" pada Label.



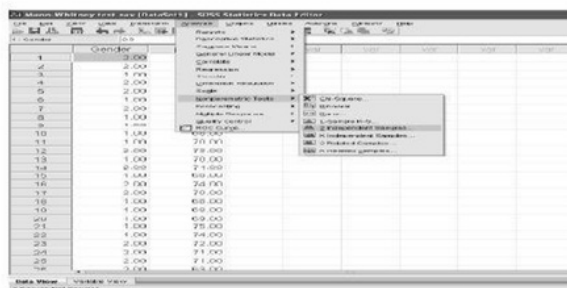
Gambar 4.22 Variabel view dengan variabel yang dimasukkan

Langkah 2 Klik pada Data view dan masukkan data seperti pada Gambar 4.23.



Gambar 4.23 Data Editor dengan data

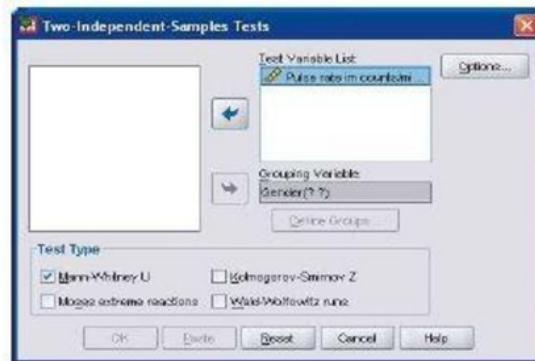
Langkah 3 Pilih Analisis, klik Non-parametric dan kemudian pilih 2 Sampel Independen (Gambar 4.24).



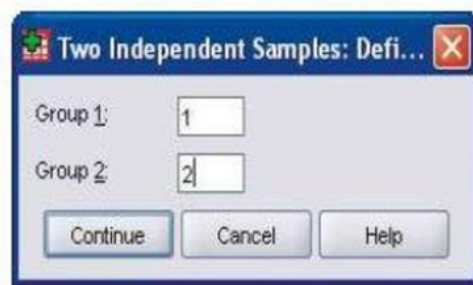
Gambar 4.24 Pilihan 2 Sampel Independen dari menu utama

Langkah 4 Dalam kotak dialog Dua-Sampel Independen (Gambar 4.23), pindahkan Pulse rate dalam hitungan/menit ke kotak **Test Variable List** dan Gender ke kotak **Grouping Variable** (langkah ini mirip dengan Uji t-sampel Independen). Sekarang klik pada Define

Groups dan ketik "1" (untuk male) di kotak Group 1 dan "2" (untuk female) di kotak Group 2 (Gambar 4.25) dan klik Continue.

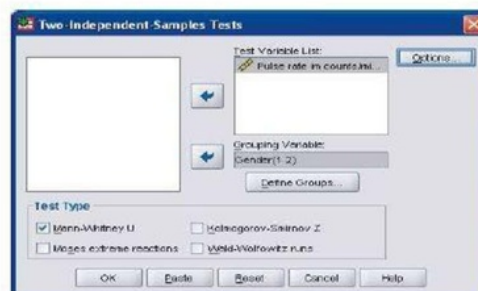


Gambar 4.25 Pilihan Mann - Whitney dari dua kotak dialog Uji Sampel Independen



Gambar 4.26 Mendefinisikan grup 1 dan 2 dalam kotak dialog Dua-Sampel Independen

Langkah 5 Langkah sebelumnya akan membawa kita kembali ke Two-Independent-Samples t-tests, sekarang pilih Mann-Whitney pada **Type Test** (Gambar 4.27).



Gambar 4.27 Uji Dua-Independen-Sampel dengan Mann-Whitney

Langkah 6 Klik OK untuk menjalankan analisis. Output muncul sebagai dua judul Peringkat dan Uji Statistik.

Output 1

Ranks				
	gender	N	Mean Rank	Sum of Ranks
Pulse rate in counts/min	male	13	12	156
	female	13	15	195
	Total	26		

Output 2

Test Statistics ^b	
	Pulse rate in counts/min
Mann-Whitney U	65.000
Wilcoxon W	156.000
Z	-1.010
Asymp. Sig. (2-tailed)	.313
Exact Sig. [2*(1-tailed Sig.)]	.336 ^a

a. Not corrected for ties.

b. Grouping Variable: gender

Interpretasi

Dalam Uji Mann-Whitney kita perlu catat uji statistik dan signifikansinya. Karena nilai Mann-Whitney adalah 65,00 dan nilai p-value untuk 2 ekor termasuk satu ekor 0,336 yang $> 0,05$, maka Hipotesis nol TIDAK DITOLAK pada p-value = 0,05. Terdapat cukup bukti untuk menyimpulkan bahwa tidak ada perbedaan dalam tingkat denyut rata-rata populasi pria dan wanita yang duduk di dalam latihan yang bergerak cepat dalam studi yang diberikan.

25

Uji Mann-Whitney untuk sampel independen (skala ordinal) dengan Uji SPSS Mann-Whitney juga dapat digunakan untuk data pada skala ordinal.

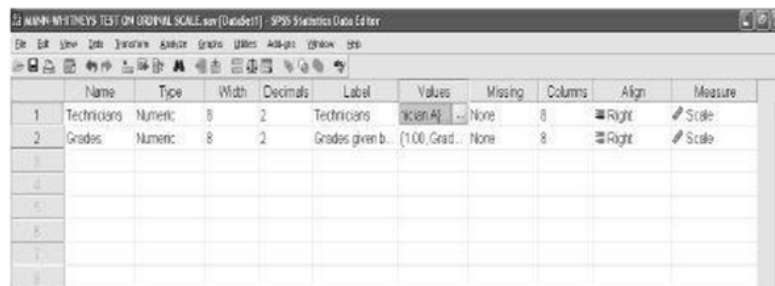
Contoh 4.6

Dua puluh lima mahasiswa sarjana dipandu oleh dua teknisi lab yang berbeda secara terpisah. Atas dasar nilai akhir, Uji Hipotesis nol bahwa mahasiswa melakukan sama baiknya di dalam kelas.

Technician A	A	A	A	A-	B	B	C+	C+	C	C	C-			
Technician B	A	A	B+	B+	B	B-	C	C	C-	D	D	D	D	D-

Hipotesis nol Kinerja mahasiswa adalah sama (setara) di bawah dua teknisi lab.

Langkah 1 Buka Data Editor, klik Variabel view, beri nama variabel sebagai "Teknisi", pilih Numerik di bawah Type, label sebagai "Teknisi" di bawah Label (Gambar 4.28). Klik pada area abu-abu di bawah Value dan ketik "1" di kotak Value dan "Teknisi A" di kotak Label dan klik Tambahkan untuk mentransfernya ke kotak di bawah dan ketik "2" di kotak Value dan "Teknisi B" di kotak Label dan klik Add untuk mentransfernya ke kotak dan klik OK. Lalu ke baris kedua dan ketik Grade di bawah Nama, "pilih Numeric" di bawah Type, ketik "Grade" di bawah Label. Di kolom "Value kolom 1 di kotak Value dan kelas A di kotak Label dan klik Add untuk mentransfernya ke kotak. Berikan Value dengan cara yang sama seperti 2, 3, 4, 5, 6, 7, 8, 9, dan 10 untuk A-, B+, B, B-, C+, C, C-, D, dan D- masing-masing dengan mengklik Add setiap kali.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Technicians	Numeric	8	2	Technicians	1=Technician A, 2=Technician B	None	8	Right	Scale
2	Grades given by...	Numeric	8	2	Grades given by...	1=A-, 2=B+, 3=B, 4=B-, 5=C+, 6=C, 7=C-, 8=D, 9=D-, 10=D	None	8	Right	Scale

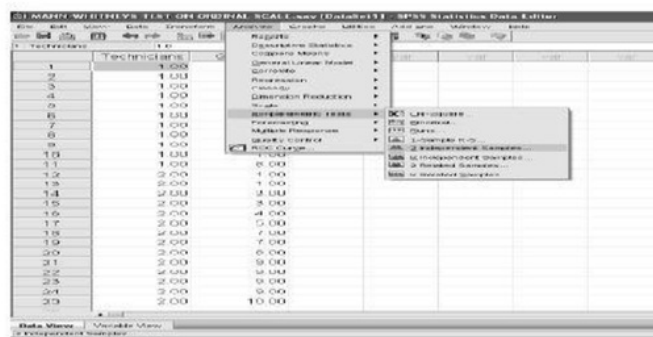
Gambar 4.28 Variable View dengan 2 variabel dimasukkan

Langkah 2 Klik pada data editor dan masukkan data di Data Editor seperti pada Gambar 4.29.

Technician	Grades					
1	1.00					
2	1.00					
3	1.00					
4	1.00					
5	1.00					
6	1.00					
7	1.00					
8	1.00					
9	1.00					
10	1.00					
11	1.00					
12	2.00					
13	2.00					
14	2.00					
15	2.00					
16	2.00					
17	2.00					
18	2.00					
19	2.00					
20	2.00					
21	2.00					
22	2.00					
23	2.00					
24	2.00					
25	2.00					
26	2.00					
27	2.00					
28	2.00					

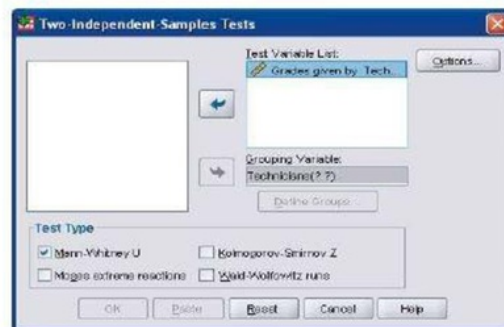
Gambar 4.29 Tampilan Data dengan data yang dimasukkan

Langkah 3 Pilih Analisis, klik Non-parametrik dan kemudian pilih 2 Sampel Independen (Gambar 4.30).



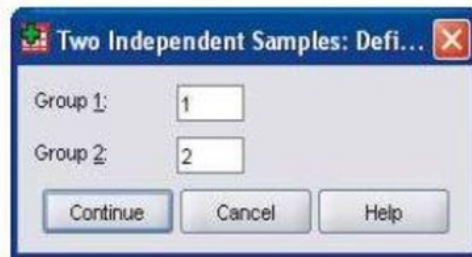
Gambar 4.30 Memilih Dua Sampel Independen dari menu utama

Langkah 4 Dalam **2-Sample Independent-Samples** (Gambar 4.31), transfer **Grade** ke kotak **Test Variable List** dan **Technician** ke kotak **Grouping Variable** (langkah ini mirip dengan Uji t-sampel Independen). Pilih Mann - Whitney U di bawah Jenis test.



Gambar 4.31 Mann - Whiney U Dipilih dalam dua kotak Uji Independent Samples

Langkah 5 Klik pada **Define Groups** dan ketik 1 (Teknisi A) di kotak Group 1 dan "2" (Teknisi B) di kotak Group 2 (Gambar 5.32) dan klik Continue.



Gambar 4.32 Menentukan kelompok dalam Dua Sampel Independen kotak dialog

Langkah 6 Klik OK untuk menjalankan analisis. Output tampak seperti yang diberikan di bawah ini.

Output 1

Ranks				
	Technicians	N	Mean Rank	Sum of Ranks
Grades given by Technicians	Technician A	11	10.14	111.50
	Technician B	14	15.25	213.50
	Total	25		

Output 2

Test Statistics ^b	
	Grades given by Technicians
Mann-Whitney U	45.500
Wilcoxon W	111.500
Z	-1.742
Asymp. Sig. (2-tailed)	.082
Exact Sig. [2*(1-tailed Sig.)]	.085 ^a

a. Not corrected for ties.

b. Grouping Variable: Technicians

Interpretasi

Output 1 memberikan jumlah kasus, rerata rangking dan jumlah rangking. Statistik Mann-Whitney U adalah 45,5 dan p-value untuk dua ekor termasuk satu ekor adalah 0,085 yang > 0,05, oleh karena itu Hipotesis nol TIDAK DITOLAK pada tingkat signifikansi 0,05. Kinerja para mahasiswa dipandu oleh dua teknisi lab adalah sama.

Uji-t Wilcoxon Matched-Pairs

Uji Wilcoxon Matched-Pairs Ranks adalah Uji non-parametrik yang sering dianggap sama dengan Uji-t berpasangan, sama seperti Uji Mann-Whitney adalah analog ke Uji-t dua sampel independen. Uji Wilcoxon signed-ranks berlaku untuk desain dua sampel yang melibatkan pengukuran berulang, pasangan yang cocok, atau langkah "sebelum" dan "sesudah". Ini berlaku untuk data pada skala rasio dan skala ordinal juga. Uji Wilcoxon digunakan untuk menentukan besar perbedaan antara kelompok yang cocok (matched).

Versi pasangan yang berpasangan (atau pasangan-sampel) (pasangan pengamatan (x_1, y_1) , (x_2, y_2) , ...) berkaitan dengan perbedaan $(x_1 - y_1)$, $(x_2 - y_2)$, ... Dengan asumsi bahwa perbedaan ini adalah pengamatan independen dari distribusi simetris, Hipotesis nol adalah bahwa distribusi ini memiliki nol median. Prosedur pengujian melibatkan perhitungan perbedaan antara setiap pasang pasangan. Kemudian satu peringkat nilai-nilai absolut dari perbedaan dari rendah ke tinggi dan kemudian memberikan tanda masing-masing perbedaan ke peringkat yang sesuai. Kemudian peringkat dengan tanda + dan peringkat dengan - diringkas secara terpisah. Untuk uji dua-ekor, kami menolak Hipotesis jika $T +$ atau $T -$ kurang dari atau sama dengan Value kritis (diberikan dalam tabel). Nilai p menjawab pertanyaan ini: Jika perbedaan rata-rata di seluruh populasi adalah nol (perlakuan tidak efektif).

Uji Wilcoxon Matched-Pairs dengan SPSS

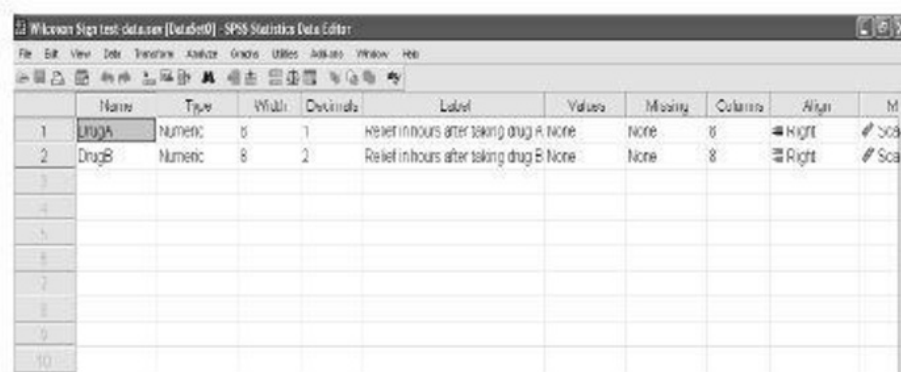
Contoh 4.7

Tabel di bawah ini menunjukkan berapa lama (jam) kesembuhan setelah meminum dua obat analgesik pada 15 pasien yang menderita arthritis. Adakah bukti bahwa satu obat memberikan kelegaan yang lebih lama daripada yang lain?

Drug A	3.0	2	3.6	2.6	7.4	3	16	7	3.3	2	6.8	8.5	7	8	5.6
Drug B	3.5	3.5	5.7	2.4	9.9	4	18.7	6.6	4.5	4	9.1	1.8	8.5	7.5	2.9

Hipotesis nol: Obat A dan Obat B memberikan kesembuhan serupa.

Langkah 1 Buka Data Editor, klik Variabel view, beri nama variabel sebagai DrugA pilih numerik di bawah Type, label sebagai Sembuh dalam jam setelah megonsumsi obat A di bawah Label, kemudian ke baris kedua dan ketik di Name sebagai DrugB pilih numerik di bawah Type, label sebagai Sembuh dalam jam setelah mengonsumsi obat B di bawah Label (Gambar 4.33).



Gambar 4.33 Variabel View dengan dua variabel

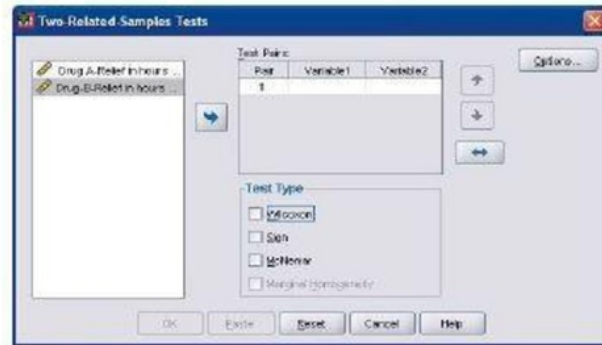
Langkah 2 Klik Tampilan Data dan masukkan data di bawah Obat A dan Obat B.

Langkah 3 Pilih Analisis, klik Non-parametrik dan pilih 2 Sampel Terkait (Gambar 4.34).



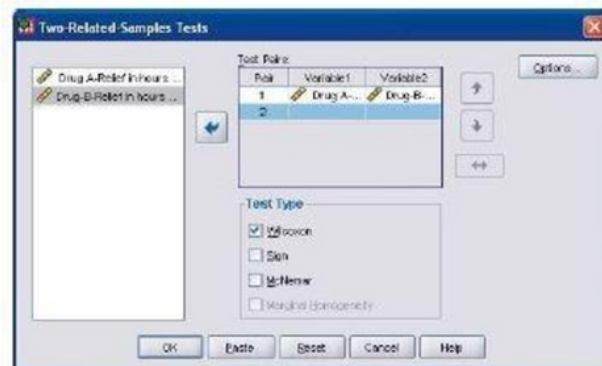
Gambar 4.34 Memilih 2 Sampel Terkait dari menu utama

Langkah 4 Sebuah jendela pop-up muncul sebagai sampel Dua-Terkait. Uji seperti yang diberikan pada Gambar 5.35.



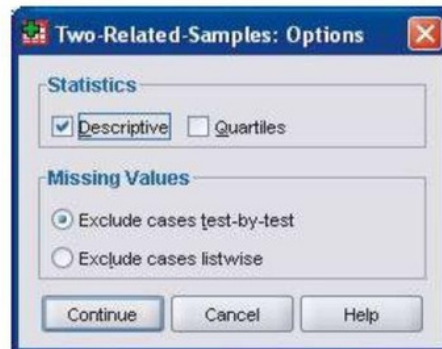
Gambar 4.35 Kotak dialog Dua Uji Sampel Terkait

Langkah 5 Transfer Obat A-Sembuh dalam jam ... ke sisi kanan bawah Variabel 1 dan Obat B-Sembuh dalam jam... di bawah Variabel 2 (Gambar 4.36).



Gambar 4.36. Variabel yang ditransfer dan Wilcoxon yang dipilih di Jenis Pengujian dalam kotak dialog Dua Sampel Terkait

Langkah 6 Pilih Wilcoxon di bawah Jenis Pengujian dan klik Opsi Two-related Samples Test (Gambar 4.39), pilih Deskriptif di Statistik dan klik Lanjutkan. Klik OK untuk menjalankan analisis.



Gambar 4.37 Deskriptif yang dipilih dalam opsi **Two-related Samples Test**

Output 1

Descriptive Statistics					
	N	Mean	Std. Deviation	Minimum	Maximum
Drug A-Relief in hours after taking drug	15	5.720	3.6688	2.0	16.0
Drug-B-Relief in hours after taking drug	15	6.1733	4.28927	1.80	18.70

Output 2

Ranks				
		N	Mean Rank	Sum of Ranks
Drug-B-Relief in hours after taking drug - Drug A-Relief in hours after taking drug	Negative Ranks	5 ^a	7.00	35.00
	Positive Ranks	10 ^b	8.50	85.00
	Ties	0 ^c		
	Total	15		

- a. Drug-B-Relief in hours after taking drug < Drug A-Relief in hours after taking drug
b. Drug-B-Relief in hours after taking drug > Drug A-Relief in hours after taking drug
c. Drug-B-Relief in hours after taking drug = Drug A-Relief in hours after taking drug

Output 3

Test Statistics ^b	
	Drug-B-Relief in hours after taking drug - Drug A-Relief in hours after taking drug
Z	-1.42 ^a
Asymp. Sig. (2-tailed)	0.16

- a. Based on negative ranks.
b. Wilcoxon Signed Ranks Test

Interpretasi

Output 1 memberikan jam rata-rata kelegaan setelah mengonsumsi obat A dan B dengan standar deviasi.

Output 2 memberikan rangking rata-rata. Uji Statistik diberikan dalam output 3. Karena Asymp. Sig. p-value (2-tailed) adalah 0,155 yang $> 0,05$, Hipotesis nol, H_0 A dan Obat B memberikan kesembuhan yang sama TIDAK DITOLAK dengan tingkat signifikansi 0,05. Oleh karena itu, tidak ada bukti bahwa satu obat memberikan kesembuhan lebih lama daripada yang lain.

REVIEW LATIHAN

1. Data berikut menunjukkan kandungan nitrat air (mg / l) dari dua danau. Analisis data dan simpulkan apakah kedua danau tersebut berbeda secara signifikan dalam kandungan nitratnya.

Lake 1	0.50	0.65	0.63	0.50	0.54	0.60	0.66	0.68	0.71	0.64
Lake 2	1.62	0.72	1.68	1.58	0.65	0.70	0.72	1.70	0.91	0.99

2. Dua atlet diuji sesuai waktu (detik) untuk menjalankan lintasan tertentu dan hasilnya diberikan di bawah ini. Ujilah apakah kedua atlet memiliki kapasitas lari yang sama?

Athletes 1	28	30	32	33	33	29	34	30	32	31
Athletes 2	29	30	32	29	28	28	30	30	31	28

3. Berikut ini adalah makanan yang dikonsumsi dalam gram oleh dua kelompok tikus. Hitung rata-rata hitung dan standar deviasi dari dua sampel yang diberikan di bawah ini. Apakah terdapat perbedaan antara dua sampel dan simpulkan preferensi satu diet di atas yang lain. Gambar diagram batang dengan error bar.

Bengal gram	65	63	67	64	68	70	71	69	75	73
Rice	58	60	61	63	62	67	66	62	68	69

4. The breath of leave (43 cm) untuk dua kelompok pohon diberikan di bawah ini. Hitung rata-rata dan standar deviasi dari dua

kelompok dan hitung signifikansi perbedaan antara dua kelompok. Berdasarkan perbedaannya apakah mungkin untuk menyimpulkan mereka berasal dari dua spesies pohon yang berbeda? Gambar diagram batang dengan error bar.

Group I	50	65	63	50	54	60	66	68	71	64
Group II	62	72	68	58	65	70	72	70	71	70

5. Kadar kolesterol (mg / 100 serum) dalam kelompok 10 individu setelah mengonsumsi obat adalah 231, 245, 208, 258, 245, 199, 252, 195, 208, dan 205. Carilah apakah obat tersebut efektif dalam mengurangi kolesterol tingkat dalam manusia. Tingkat kolesterol normal pada manusia adalah 190mg / 100 serum.

6. Obat anti-depresif diberikan kepada 8 pasien dan tekanan darah (mm Hg) sebelum dan sesudah pemberian obat diberikan di bawah ini. Temukan efek obat dalam mengurangi tekanan darah.

Before drug administration	145	135	135	136	150	131	136	154
After drug administration	123	130	134	127	123	124	119	118

7. Sebuah perusahaan farmasi mengembangkan obat, yang diklaim untuk meningkatkan kadar hemoglobin pada orang tua. Kandungan hemoglobin (g/100 ml) dari sepuluh subyek diukur sebelum dan sesudah pemberian obat. Tentukan apakah klaim perusahaan valid setelah mengamati data berikut.

Subject	1	2	3	4	5	6	7	8	9	10
Before	10	9	11	12	8	7	12	18	10	9
After	12	11	13	14	9	10	12	14	11	12

Bab 5

ANALISIS VARIANS (ANOVA)

Pada Bab 4, pengukuran variabel diperoleh untuk dua sampel dan satu sampel Uji-t atau setara non-parametrik Uji-t digunakan dalam pengujian hipotesis dan menurunkan kesimpulan. Seringkali peneliti memperoleh pengukuran variabel pada tiga atau lebih sampel dari tiga atau lebih populasi. Jika kita menerapkan Uji dua sampel dan menyatakan hipotesis nol, itu tidak akan sesuai. Melakukan serangkaian Uji dua sampel untuk pengujian hipotesis tidak valid karena alasan berikut. Ketika kita melakukan setiap Uji dua sampel pada tingkat signifikansi 5% (keyakinan), ada 95% kemungkinan bahwa kita akan tidak menolak H_0 ketika dua populasi berarti sama. Untuk satu set tiga hipotesis yang diajukan di atas, kemungkinan untuk menerima semuanya hanya $0,953 = 0,86$. Probabilitas dengan salah menolak setidaknya satu dari hipotesis nol adalah $1 - 0,86 = 0,14$. Tingkat signifikansi menjadi 0,14 bukan 0,05. Untuk alasan ini, kita tidak memilih Uji-t dalam pengujian hipotesis untuk lebih dari dua rata-rata; sebagai gantinya kita memilih prosedur yang menguji ekuivalensi rata-rata. Prosedur yang tepat adalah faktor tunggal analisis varians, sering disingkat sebagai ANOVA (Analysis of Variance), yaitu F-Test. Di sini, Hipotesis nol yang tepat adalah $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$. Jika ada perbedaan yang signifikan, ada ketentuan untuk Uji perbandingan post hoc-multiple. Uji Post hoc-multiple comparison memungkinkan kita untuk menemukan range sebagai jawaban dan menguji signifikansi perbedaan antara rata-rata hitung juga dalam berpasangan.

ANALISIS VARIANS — SATU FAKTOR ANTARA SUBYEK (ANOVA)

Analisis Varians, disingkat sebagai ANOVA, dikembangkan oleh R.A. Fisher; sebenarnya Uji-F dinamai untuk menghormatinya. R.A. Fisher menekankan pentingnya keacakan, yaitu, ukuran sampel identik tidak diperlukan untuk faktor tunggal ANOVA, tetapi ukuran sampel harus sedapat mungkin sama. Faktor tunggal ANOVA dikatakan mewakili desain eksperimental yang sepenuhnya acak (**completely randomized experimental design**). Dalam ANOVA kita mengasumsikan bahwa $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$ dan memperkirakan varians populasi dan varians digolongkan sebagai total varians, antara varians (between variance) dan dalam varians (within variance). Dengan demikian, ANOVA didasarkan pada porsi variasi dalam

variabel dependen. Ini membandingkan varians antara kelompok dengan varians dalam kelompok. Jika ada lebih banyak perbedaan antara kelompok daripada kelompok, maka kelompoklah yang membuat perbedaan dan hasilnya secara statistik signifikan.

$$F = \frac{MS_{\text{between}}}{MS_{\text{within}}}$$

MSbetween – rata-rata jumlah kuadrat antara sampel,
MSwithin – rata-rata jumlah kuadrat dalam sampel.

Semakin besar nilai F-ratio, semakin besar varians antar kelompok.

ANOVA SATU ARAH (ONE-WAY)

Jika efeknya hanya diuji untuk satu faktor (satu variabel independen) pada variabel (dependen) yang bersangkutan, maka itu disebut analisis univariat.

Catatan. Ukuran sampel yang identik tidak diperlukan untuk ANOVA Satu faktor, tetapi ukuran sampel sedapat mungkin hampir sama. Tidak ada aturan tegas untuk jumlah pengamatan yang diperlukan. Ini dapat diputuskan oleh peneliti di lapangan.

Dasar dan makna One-way ANOVA dilakukan hanya untuk satu variabel independen dan sampel milik kelompok yang berbeda dari populasi yang sama. Jika peneliti ingin menguji apakah empat pakan yang berbeda menghasilkan berat badan yang berbeda pada unggas, “pakan” adalah faktor dan “berat badan” adalah variabel. Berbagai jenis pakan dikatakan tingkat faktor. Ketika peneliti mendesain eksperimennya, hewan percobaan diambil secara acak untuk menerima satu dari empat pakan dengan jumlah hewan yang kurang lebih sama yang menerima setiap pakan. Data yang diberikan di bawah ini dengan perhitungan manual F-value dengan menjelaskan dasar-dasar ANOVA satu arah.

A	B	C
2	9	10
3	6	6
1	8	9
5	7	7
4	5	8
$K_A = 5$	$K_B = 5$	$K_C = 5$

Hipotesis nol: ¹¹ Tidak ada perbedaan yang signifikan antara kelompok yang berbeda.

Rata-rata hitung keseluruhan = 6

$$N=15, \sum X_A=15, \sum X_B=35, \sum X_C=40, \sum X^2=640, \sum X=90$$

Rata-rata A = 3 Rata-rata B = 7 Rata-rata C = 8

Langkah 1 Hitung jumlah total kuadrat. Harap diingat bahwa "jumlah kuadrat" mengacu pada jumlah deviasi kuadrat dari semua anggota distribusi dari rata-rata distribusi itu. $SS_{\text{total}} = \sum (X - \bar{X})^2$

$$= (2 - 6)^2 + (3 - 6)^2 + (1 - 6)^2 + \dots + (8 - 6)^2$$

$$= 100$$

Langkah 2 Hitung *SS between groups* Istilah "antar kelompok" secara tradisional digunakan. Proses ini melibatkan penjumlahan deviasi kuadrat antara setiap rata-rata hitung kelompok dan rata-rata hitung besar yang memastikan bahwa kita mengalikan setiap deviasi kuadrat dengan jumlah (K) dalam setiap kelompok.

Demikian,

$$\begin{aligned} SS_{\text{between}} &= K_A (X - \bar{X}_A)^2 + K_B (X - \bar{X}_B)^2 + K_C (X - \bar{X}_C)^2 \\ &= 5(6-3)^2 + 5(6-7)^2 + 5(6-8)^2 \\ &= 45 + 5 + 20 = 70 \end{aligned}$$

Langkah 3 Hitung dalam-kelompok (error) *SS Pertama*, jumlahkan penyimpangan kuadrat antara setiap pengukuran dalam kelompok tertentu dan kemudian jumlahkan dalam kuadrat dari ketiga kelompok.

Jadi,

A	$(X_A - \bar{X}_A)^2$	B	$(X_B - \bar{X}_B)^2$	C	$(X_C - \bar{X}_C)^2$
2	1	9	4	10	4
3	0	6	1	6	4
1	4	8	1	9	1
5	4	7	0	7	1
4	1	5	4	8	0
Total	10		10		10

$$SS_{\text{within}} = 10 + 10 + 10 = 30$$

Kita sekarang telah membagi total variabilitas data menjadi komponen "antar-kelompok" dan komponen "dalam-kelompok," atau kesalahan. Langkah selanjutnya adalah menghitung dua varians yang terlibat dan membandingkannya, untuk itu tabel ANOVA diformulasikan.

Langkah 4 Merumuskan tabel ANOVA

ANOVA table			
Source	Degrees of freedom	Sum of squares	Mean square
Between	2	70	35.00
Within (error)	12	30	2.50
Total	14	100	

Derajat kebebasannya adalah $N - 1$. Karena ada total 15 variasi, derajat kebebasan totalnya adalah $15 - 1 = 14$ untuk total. Jumlah derajat kebebasan antar kelompok diperoleh dengan mengurangi 1 dari total jumlah kelompok; karena ada 3 kelompok, jumlah derajat kebebasan untuk antar kelompok adalah $3 - 1 = 2$. Dalam kelompok derajat kebebasan = $15 - 3 = 12$. Ada tiga kelompok, masing-masing dengan 5 variasi; oleh karena itu $(5-1) + (5-1) + (5-1) = 12$.

Langkah 5 Hitung rata-rata jumlah kuadrat (MS) $MS_{\text{between}} = 70/2 = 35,00$

$MS_{\text{within}} = 30/12 = 2.50$

Langkah 6 Hitung F-ratio

Langkah terakhir dalam analisis varians melibatkan perbandingan antara-kelompok varians dengan varians dalam-kelompok untuk menentukan apakah antara-kelompok varians secara signifikan lebih besar. Dengan kata lain, kita ingin tahu apakah efek pengobatan signifikan atau apakah perbedaan antar kelompok dapat dijelaskan hanya berdasarkan variasi acak.

$$\begin{aligned}
 F_{2,12} &= \frac{MS_{\text{between}}}{MS_{\text{within}}} \\
 &= \frac{35}{2.50} = 14.00
 \end{aligned}$$

Interpretasi:

Kita membaca F-value untuk 2 dan 12 df dari F-tabel (kita turun ke kolom kedua ke baris kedua belas). 2 derajat kebebasan dikaitkan dengan pembilang dan 12 derajat kebebasan dengan

penyebutnya. Di mana kolom dan baris berpotongan, kita menemukan 3,88 sebagai Value F kritis untuk level 0,05 dan 6,93 sebagai nilai kritis untuk level 0,01. Karena nilai F yang dihitung adalah 14,00, yang lebih besar dari nilai F tabel pada 0,01 tingkat signifikansi. Kita MENOLAK Hipotesis nol. Oleh karena itu, kita dapat menyimpulkan bahwa efek pengobatan telah menghasilkan perbedaan yang signifikan secara statistik di antara kelompok-kelompok tersebut. Dalam analisis satu faktor, ketika ada perbedaan yang signifikan dalam varians antara sampel dimana Hipotesis nol ditolak dan uji perbandingan ganda dilakukan untuk menentukan perbedaan nilai rata-rata hitung sampel yang terjadi. Dalam analisis satu faktor, jika F-value tidak signifikan, tidak perlu melakukan uji perbandingan Post hoc-multiple.

Post hoc multiple comparison Dalam melakukan analisis satu faktor varians, kita menguji hipotesis nol $H_0: \mu_1 = \mu_2 = \mu_3 = \dots \mu_4$ dan seterusnya. Namun, ketika kita menolak Hipotesis nol, itu tidak berarti bahwa semua rata-rata populasi ($\mu_1, \mu_2, \mu_3, \mu_4$, dll.) berbeda satu sama lain. Lebih jauh lagi, kita tidak tahu berapa banyak rata-rata yang berbeda satu sama lain dan di mana perbedaannya terletak di antara jumlah rata-rata populasi yang diberikan. Masalah ini diatasi oleh beberapa Uji perbandingan. Beberapa Uji perbandingan diinginkan untuk analisis satu faktor (Model I ANOVA dan bukan untuk Model II ANOVA). Secara umum, beberapa Uji perbandingan untuk rata-rata memiliki asumsi dasar yang sama seperti analisis varians yaitu populasi terdistribusi secara normal dan variasinya homogen. Dalam semua pengujian perbandingan ganda, ukuran sampel yang harus sama, tetapi kadang-kadang dilakukan dengan sampel yang tidak sama.

Ada sebanyak 18 uji post hoc-multiple comparison. Sebagian besar uji dirancang untuk memeriksa perbedaan pasangan. Mereka memberikan informasi tambahan tentang variabel yang sama. Pilihan perbandingan bergantung pada situasi yang sama yang dimiliki seseorang.

Kesimpulan pada beberapa pengujian perbandingan tergantung pada pertimbangan urutan perbandingan pasangan. Prosedur yang benar adalah untuk pertama membandingkan rata-rata hitung terbesar melawan yang terkecil, kemudian yang terbesar melawan yang lebih kecil berikutnya dan seterusnya, hingga yang terbesar dapat dibandingkan dengan yang terbesar kedua. Lalu terbesar kesatu membandingkan yang terbesar kedua dengan yang terkecil, terbesar kedua dengan terkecil berikutnya dan seterusnya.

Ada sejumlah uji perbandingan ganda, namun tidak ada kesepakatan mengenai prosedur terbaik untuk digunakan secara rutin. Uji yang paling banyak dan umum digunakan adalah Uji Tukey, Uji Newman-Keuls dan Uji Duncan, sering disebut sebagai "Duncan Multiple Range Test".

Kadang-kadang beberapa uji perbandingan akan menghasilkan kesimpulan yang ambigu dalam bentuk kesamaan yang tumpang tindih. Dalam beberapa kasus, misalnya, sampel 1 dan 2 membentuk subset tunggal yang menunjukkan bahwa kedua sampel berasal dari populasi 1 dan sampel 2, 3 dan 4 membentuk satu bagian yang menunjukkan bahwa ketiga sampel ini berasal dari populasi 2. Dalam kasus ini, sampel 2 ditugaskan untuk populasi 1 dan populasi 2 yang tidak mungkin. Dengan demikian kita hanya dapat menyatakan itu tetapi kita tidak dapat menyimpulkan bagaimana hal itu terkait dengan dan dalam situasi ini, menaikkan ukuran sampel (jumlah data yang lebih besar) akan memberikan kesimpulan yang tepat.

Jika ukuran sampel lebih besar, maka hasil dari beberapa uji perbandingan akan mencari perbedaan antar rata-rata. Salah satu batasan dari beberapa uji perbandingan adalah ketidakmampuannya untuk menentukan posisi beberapa rata-rata secara akurat.

Tukey test multiple comparison test Uji ini adalah Uji perbandingan yang banyak digunakan. Ini terdiri dari Hipotesis nol versus Hipotesis alternatif. Misalnya, jika ada 4 grup (1, 2, 3, dan 4). Uji Tukey membandingkan 1 dan 2, 1 dan 3, 1 dan 4 dan kemudian 2 dan 1, 2 dan 3 dan 2 dan 4 dan seterusnya.

Uji Tukey, juga dikenal sebagai Uji rentang Tukey, metode Tukey, Uji signifikansi Tukey sesungguhnya, Uji Tukey HSD (Honestly Significant Difference, atau metode Tukey-Kramer, merupakan prosedur perbandingan beberapa langkah dan Uji statistik yang umumnya digunakan bersama dengan ANOVA untuk menemukan cara yang berbeda secara signifikan satu sama lain. Dinamakan setelah Tukey, membandingkan semua kemungkinan pasangan rata-rata dan didasarkan pada distribusi rentang mahasiswa (distribusi ini mirip dengan Distribusi-t. Uji ini membandingkan rata-rata setiap perlakuan dengan rata-rata setiap perlakuan lainnya, yaitu, berlaku secara bersamaan untuk himpunan semua perbandingan berpasangan dan mengidentifikasi dimana perbedaan antara dua rata-rata lebih besar.

Duncan multiple range test (MRT) Uji rentang ganda Duncan (MRT) adalah prosedur perbandingan berganda yang dikembangkan oleh David B. Duncan pada tahun 1955. MRT Duncan termasuk kelas umum

prosedur perbandingan berganda yang menggunakan statistik rentang mahasiswa untuk membandingkan set rata-rata. Uji ini adalah varian dari metode Student - Newman - Keuls yang menggunakan peningkatan tingkat alfa untuk menghitung nilai kritis di setiap langkah. Ini sangat protektif terhadap False Negative (Tipe II) dengan mengorbankan memiliki risiko lebih besar membuat False positive (Type I). Ini biasanya digunakan dalam agronomi dan penelitian pertanian lainnya. Uji Duncan telah dikritik karena terlalu liberal oleh banyak ahli statistik termasuk Henry Scheffé dan John W. Tukey. Duncan berpendapat bahwa prosedur yang lebih liberal adalah tepat karena di dunia nyata mempraktekkan secara global Hipotesis nol.

H_0 = "Semua cara adalah sama" sering salah dan dengan demikian para ahli statistik tradisional terlalu melindungi Hipotesis nol palsu terhadap kesalahan tipe I. Jika peneliti ingin menemukan kisaran dalam tanggapan dari kelompok perlakuan yang berbeda untuk pengobatan tertentu atau respon homogen dalam kelompok perlakuan yang berbeda maka Uji rentang berganda Duncan adalah yang paling sesuai.

Contoh 5.1

Data berikut kadar gula darah (mg/100 ml) diperoleh dari laboratorium klinis. Lakukan analisis varians antara kelompok dan hitung efektivitas herbal pada kadar gula darah.

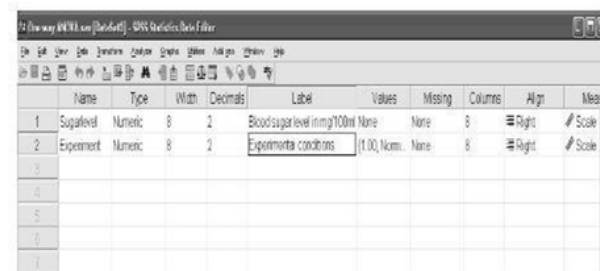
Normal	Diabetic	Herb 1	Herb 2
96.00	180.00	180.00	120.00
100.00	225.00	190.00	130.00
111.00	260.00	185.00	130.00
98.00	250.00	190.00	135.00
106.00	265.00	180.00	136.00
105.00	280.00	170.00	140.00

Hipotesis nol: Cara kelompok sama dengan satu sama lain.

ANOVA satu arah dengan SPSS

Langkah 1 Buka Data Editor, klik Variabel view, (sebelum mulai memasukkan data, tentukan "factor" dan "level" dan variabel yang akan diuji, dalam hal ini empat kelompok adalah tingkat faktor

yang sama, yaitu perlakuan (eksperimen) dan tingkat Gula adalah variabel Name variabel sebagai "Sugarlevel" di bawah Name, pilih Numeric pada Type, beri Label sebagai "Blood Sugar level dalam mg/100 ml" (Gambar 5.1).



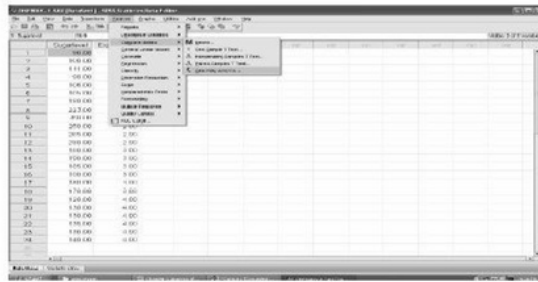
Gambar 5.1 Penampilan Variabel view dengan rincian 2 variabel

Langkah 2 Lanjut ke nama baris kedua variabel sebagai "Experiment" pada Name, pilih Numeric pada Type, beri label sebagai "Experimental Conditions" pada Label. Klik kolom Value dan klik dua kali pada area abu-abu di bawah Value. Jendela popup terbuka seperti pada Gambar 5.2, ketik "1" di kotak Value dan Normal di kotak Label dan klik Add untuk memindahkannya ke kotak bawah. Ketik "2" di kotak Value dan "Diabetes" di kotak Label dan klik Add untuk memindahkannya ke bawah, lanjutkan dengan cara yang sama untuk sisa data sebagai "3" untuk "Herb" 1 "dan" 4 "Untuk" Herb "2" "dengan menambahkan setiap waktu dan klik OK.



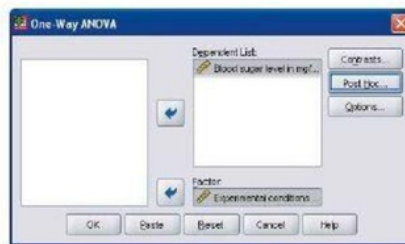
Gambar 5.2 Kotak dialog Value Labels dengan kode variabel

Langkah 3 Ketik data dalam Data view di bawah setiap judul. Pilih Analyze dari menu pilih Compare Rata-rata Means, lalu klik One-way ANOVA (Gambar 5.3).



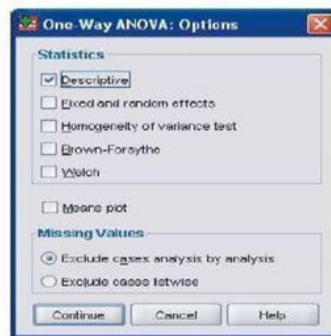
Gambar 5.3 Memilih ANOVA Satu Arah dari menu utama

Langkah 4 1ptak dialog ANOVA satu arah terbuka (Gambar 5.4). Pindahkan (Blood Sugar level dalam mg/100 ml) ke **Dependent List** dan **Experimental Conditions** ke **factor** dengan mengklik panah.



Gambar 5.4 Kotak dialog ANOVA satu arah dengan variabel dan Faktor ditransfer

Langkah 5 Klik Opsi. Jendela popup terbuka arahkan ke One-Way ANOVA: Opsi (Gambar 5.5), pilih Deskriptif dan klik Lanjutkan dan klik OK untuk menjalankan analisis.



Gambar 5.5 One-way ANOVA: Pilihan untuk memilih Deskriptif
Output muncul seperti pada **Output 1** dan **Output 2**.

Output 1

Blood sugar level in mg/100ml							
Descriptives							
	N	Mean	Std. Deviation	Std. Error	90% Confidence Interval for Mean		
					Lower Bound	Upper Bound	
Normal	6	102.6667	5.64506	2.32458	96.7425	108.5908	96.00
Diabetic	6	243.3333	35.09326	14.70772	205.5439	281.1227	180.00
Herb 1	6	162.5000	7.28269	3.09279	154.5429	170.4571	170.00
Herb 2	6	131.8333	6.94322	2.83333	126.5500	139.1166	120.00
Total	24	165.0833	57.38422	11.71351	140.8521	189.3145	96.00

Output 2

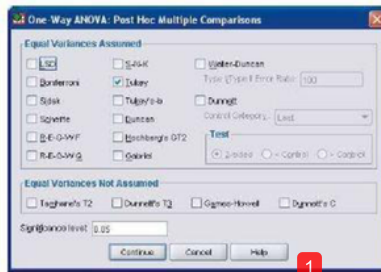
ANOVA					
Blood sugar level in mg/100ml					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	68566.833	3	22855.611	63.745	.000
Within Groups	71711.000	20	358.550		
Total	75737.833	23			

Interpretasi

Rata-rata, standar deviasi dan standar deviasi dari empat sampel diberikan dalam Output 1 di bawah judul Deskriptif. Hasil ANOVA diberikan dalam Output 2. Signifikansi varians, yaitu, p-value, yang diberikan⁵ di bawah judul Sig. adalah 0,000. Karena, Value signifikansi 0,000 kurang dari 0,05 ($p < 0,05$), varians¹⁹ antara kondisi percobaan yang berbeda adalah signifikan. Oleh karena itu, kita menyimpulkan bahwa empat sampel rata-rata berbeda satu sama lain secara signifikan.

Jika kita hentikan interpretasi sampai disini mungkin kita tidak mengerti apakah perbedaan antara varians sampel adalah sama atau dua varians sampel adalah sama. Selain itu, kita mungkin tidak menemukan sumber perbedaan signifikan di antara rata-rata kelompok. Sebagian besar Uji post hoc dirancang untuk memeriksa perbedaan pasangannya. Pilihan perbandingan tergantung pada situasi yang Anda miliki. Perbandingan ganda yang tidak terencana di antara level faktor dapat dilakukan dengan mengklik salah satu opsi dalam beberapa perbandingan post-hoc, tetapi peneliti harus memiliki ide tentang data dan pengetahuannya pada Uji post-hoc dengan baik, hanya dengan itu peneliti tersebut akan dapat menambang fakta menarik tentang datanya. Dalam hal ini karena ukuran sampel sama, kita yakin memiliki varians populasi yang sama. Karena kita membutuhkan perbandingan dan rentang pasangannya juga, kita lanjut untuk beberapa perbandingan post-hoc dan memilih Tukey.

Langkah 6 Klik Post-Hoc jendela popup terbuka (Gambar 5.6), dan pilih Tukey dan klik Lanjutkan dan klik OK.



Gambar 5.6 Uji Tukey dipilih dalam One-way ANOVA: Post Hoc Multiple Comparisons

Langkah 7 Sekarang **One-Way ANOVA**: Kotak dialog **Post-Hoc Multiple Comparison** terbuka. Pilih Tukey lalu klik Lanjutkan. Terakhir klik OK untuk menjalankan analisis. Output untuk muncul di layar.

Langkah 8 Simpan output (Output 3 dan 4) dan tafsirkan hasilnya.

Output 3

Multiple Comparisons

Based upon the adjusted Tukey HSD

		Mean Difference (I - J)			95% Confidence Interval	
I: Experimental condition		J: Reference condition		Sig.	Lower Bound	Upper Bound
Normal	Diabetic	-140.66667	10.93237	.000	-171.2657	-110.0677
	Herb 1	-70.83333	10.93237	.000	-110.4323	-49.2345
	Herb 2	-25.16667	10.93237	.065	-59.7657	1.4323
Diabetic	Normal	140.66667	10.93237	.000	110.0677	171.2657
	Herb 1	60.83333	10.93237	.000	20.2345	81.4323
	Herb 2	111.50000	10.93237	.000	80.9010	142.0990
Herb 1	Normal	70.83333	10.93237	.000	49.2345	110.4323
	Diabetic	-60.83333	10.93237	.000	-81.4323	-49.2345
	Herb 2	90.66667	10.93237	.001	20.6677	81.2657
Herb 2	Normal	25.16667	10.93237	.045	-1.4323	51.7657
	Diabetic	-111.50000	10.93237	.000	-142.0990	-80.9010
	Herb 1	-50.66667	10.93237	.001	-81.2657	-20.6677

*. The mean difference is significant at the 0.05 level.

Interpretasi

Pada Output 3, di baris pertama, **Normal** dibandingkan pasangan dengan tiga kelompok lainnya. Dari sini kita dapat menyimpulkan bahwa gula darah dalam Normal berbeda secara signifikan dari Diabetes dan Herbal 1 tetapi Normal tidak berbeda secara signifikan dari Herbal 2 ($p\text{-value} = 0,065$). Kesimpulan ilmiah adalah bahwa pasien yang diobati dengan Herbal 2 memiliki tingkat gula darah yang mirip dengan orang normal, karena itu Herbal 2 dapat digunakan sebagai agen antidiabetes ($p > 0,05$). Hasil yang diberikan di baris kedua membandingkan tingkat gula darah diabetes dengan tiga kelompok lainnya. Karena semua Value di bawah Sig. kolom adalah 0,000, tingkat gula darah diabetes berbeda secara signifikan dari tiga kelompok lainnya ($p < 0,01$). Baris ketiga membandingkan Herbal 1 dengan tiga kelompok lainnya secara berpasangan. Di sini juga ada perbedaan yang signifikan.

Yang keempat baris menyimpulkan yang sama yang dinyatakan dalam tiga baris sebelumnya dalam perbandingan berpasangan.

Output 4 Subset homogen

Blood sugar level in mg/100ml

Tukey HSD^a

Experimental conditions	N	Subset for alpha = 0.05		
		1	2	3
Normal	6	102.6667		
Herb 2	6	131.8333		
Herb1	6		182.5000	
Diabetic	6			243.3333
Sig.		.065	1.000	1.000

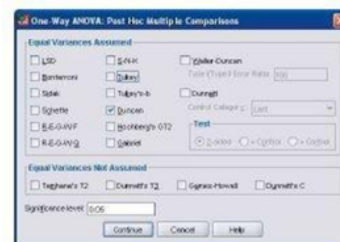
Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 6.000.

Interpretasi

Output 4 memberikan subset Homogen. Dalam Subset 1 ini dibentuk oleh **Normal** dan Herbal 2, Subset 2 diwakili oleh Herbal 1 dan Subset 3 oleh diabetes. Dari sini kita dapat menyimpulkan bahwa tingkat gula darah dalam kelompok normal adalah menuju ujung bawah kisaran gula darah, yang diikuti oleh Herbal 1 dan Diabetes adalah menuju ujung yang lebih tinggi. Karena, Normal dan Herbal 1 membentuk satu bagian, pengobatan dengan Herbal 2 akan menurunkan kadar gula darah dan bukan Herbal 1.

Jika peneliti tidak tertarik dalam perbandingan berpasangan dan jika perlu menemukan hanya rentang tanggapan maka dapat menjalankan Uji Post-hoc Duncan.

Langkah 9 Dalam One-Way ANOVA: Perbandingan multiple post-hoc, pilih Duncan dan klik Lanjutkan dan kemudian klik OK. Output tampak seperti pada Output 5.



Gambar 5.7 One-way ANOVA: Post Hoc Multiple Comparisons dengan pilihan Duncan

Output 5

Blood sugar level in mg/100ml

Duncan^a

Experimental conditions	N	Subset for alpha = 0.05			
		1	2	3	4
Normal	6	102.6667			
Herb 2	6		131.8333		
Herb1	6			182.5000	
Diabetic	6				243.3333
Sig.		1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 6.000.

Interpretasi

Output 5 memberikan hasil Uji perbandingan berganda Duncan. Di sini hasilnya dapat dikategorikan ke dalam empat himpunan bagian yang menunjukkan bahwa tidak ada satu pun dari keempat kelompok itu yang serupa satu sama lain. Kadar gula darah kelompok Herbal 2 lebih dekat ke normal, sedangkan kelompok diabetes jauh dari normal. Oleh karena itu, Uji ini berguna dalam mengkategorikan cara kelompok yang berbeda dari Value yang lebih rendah ke yang lebih tinggi dan menemukan kedekatan kelompok.

ANOVA DUA ARAH (TWO WAY ANOVA) – EKSPERIMEN DAN ANALISIS DUA FAKTOR PERLAKUAN

Ini adalah prosedur untuk menganalisis perbedaan rata-rata dari tiga atau lebih kelompok rata-rata variabel dependen yang dipengaruhi oleh lebih dari satu faktor independen. Desain penelitian terdiri dari menghitung efek dari dua variabel independen pada satu variabel dependen pada saat yang sama, yaitu, efek simultan dari kedua variabel independen pada variabel dependen. Ini menunjukkan efek keseluruhan dari dua variabel independen dan juga apakah ada interaksi diantara variabel tersebut.

ANOVA DUA ARAH DENGAN SPSS

Contoh 5.2

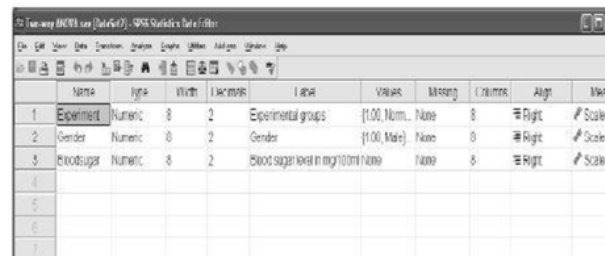
Kerjakan dengan ANOVA dua arah untuk menemukan pengaruh interaktif herbal dan seks pada tingkat gula darah (mg/100 ml).

Normal		Diabetic		Herb 1		Herb 2	
Male	Female	Male	Female	Male	Female	Male	Female
96	101	180	270	180	192	120	130
100	97	225	250	190	170	130	125
111	110	260	230	185	169	130	126
98	108	250	220	190	198	135	140
106	102	265	267	180	174	136	138
105	111	280	284	170	180	140	119

Langkah 1 Buka Data Editor, klik Variabel view (sebelum mulai memasukkan data, tentukan faktor dan tingkat dan variabel yang akan diuji). Dalam hal ini, empat kelompok adalah tingkat dua faktor yaitu Eksperimen dan Gender. (Level Gula Darah adalah variabel), di kolom pertama, ketik "Experiment", pilih Numerik pada Type, beri label sebagai "Experimental Condition" pada Label (Gambar 5.8). Pada kolom Value ketik "Experimental Group", dan klik pada area abu-abu, jendela popup terbuka (detail dapat dimasukkan seperti pada contoh sebelumnya), ketik "1" di kotak Value dan "Normal" di kotak Label dan klik Add untuk pindahkan ke kotak bawah. Ketik "2" di kotak Value dan "Diabetic" di Label kotak dan klik Add untuk mentransfer ke kotak ke bawah dan lanjutkan dengan cara yang sama untuk sisa data sebagai "3" untuk Herb 1 dan "4" untuk Herb 2 dengan menambahkan setiap waktu dan klik OK.

Langkah 2 Di baris kedua, ketik "Gender" pada Name, pilih "Numerik" pada Type, beri label "Gender" di Label (Gambar 5.7). Di kolom Value, ketik "Gender", dan klik pada area abu-abu, di jendela munculan ketik "1" di kotak Value dan Male di kotak Label dan klik Add untuk ditransfer ke kotak bawah. Ketik "2" di kotak Value "Female" di Label, klik Add dan klik OK.

Langkah 3 Pada baris ketiga, ketik "Bloodsugar" pada Name, pilih Numeric pada Type, beri label "Blood Sugar Level mg/100ml" di Label.

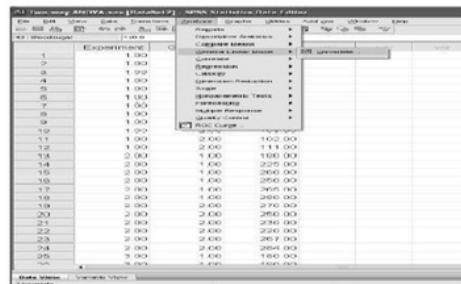


	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Experiment	Numeric	8	2	Experimental groups	{100, None, None}	8		Right	Scale
2	Gender	Numeric	8	2	Gender	{100, Male, None}	8		Right	Scale
3	Bloodsugar	Numeric	8	2	Blood sugar level in mg/100ml	None	8		Right	Scale
4										
5										
6										
7										

Gambar 5.8 Variabel view dengan rincian tiga variabel

Langkah 4 Ketik data dalam Data view di setiap judul (Gambar 5.8). Harap hati-hati untuk memasukkan data sedemikian rupa sehingga empat kategori di kolom pertama dan di kolom kedua 1 (Male) dan 2 (Female) terhadap setiap kelompok eksperimen dan masukkan level gula darah di kolom ketiga sesuai dengan eksperimen dan gender.

Langkah 5 Pilih Analyze dari menu pull-down, pilih **General Linear Model** dan kemudian klik **Univariate**.



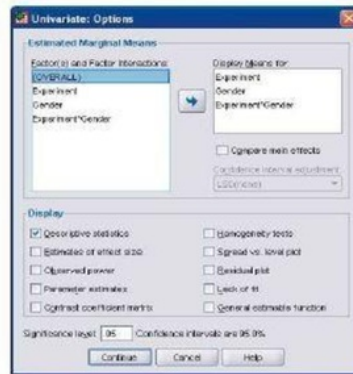
Gambar 5.9 Memilih Univariate dari menu utama

Langkah 6 Jendela popup univariat terbuka seperti pada Gambar 6.9. Pindahkan **Blood Sugar Level mg/100 ml** ke kotak Variabel Dependen, dan kelompok Eksperimental dan Gender ke **Fixed Factor(s)**.



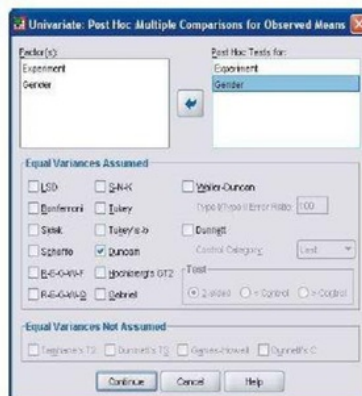
Gambar 5.10 Kotak Univariat untuk mentransfer variabel

Langkah 7 Klik Opsi untuk mendapatkan Univariate: Opsi kotak dan pindahkan Eksperimen, Gender, dan Experiment*Gender ke kotak **Display Means for:** kotak (Gambar 5.10). Pilih Statistik deskriptif di tampilan dan klik Lanjutkan.



Gambar 5.11 Univariate: Pilihan kotak Display Means for dan statistik Deskriptif

Langkah 8 Klik **Post-Hoc**, jendela popup (Univariate: Post-Hoc Multiple Comparison for Observed Means) terbuka (Gambar 5.12). Pindahkan Eksperimen dan Gender ke Uji Post-Hoc : kotak. Pilih Duncan pada **Equal Variances Assumed** dan klik Continue dan klik OK untuk menjalankan analisis.



Gambar 5.12 Univariate: Post Hoc Comparisons for Observed Means dengan variabel Uji Post Hoc dan Duncan dipilih

Output muncul di judul yang berbeda; Anda mungkin tidak membutuhkan semua output. Pilih Statistik Deskriptif (Output 1), **Tests of between Subject Effects** (Output 2), **Experimental*Gender** (Output 3) dan Uji rentang berganda Duncan dengan kadar gula Darah dalam mg/100 ml (Output 4). Dalam contoh ini karena Gender hanya memiliki dua kategori, uji post-hoc tidak dapat dilakukan (untuk melakukan Uji rentang ganda kita membutuhkan setidaknya tiga kategori).

Output 1

Descriptive Statistics

Dependent Variable: Blood sugar level in mg/100ml

Experimental groups	Gender	Mean	Std. Deviation	N
Normal	Male	102,6667	8,64008	6
	Female	104,8333	8,63819	6
	Total	103,7500	8,63957	12
Diabetic	Male	243,3333	36,00000	6
	Female	263,6000	24,78608	6
	Total	248,4167	28,94880	12
Herb 1	Male	182,0000	7,88288	6
	Female	180,5000	12,02913	6
	Total	181,5000	9,84305	12
Herb 2	Male	131,8333	8,94022	6
	Female	128,6667	8,06859	6
	Total	130,7500	7,26292	12
Total	Male	168,0833	87,38422	24
	Female	167,1250	68,84384	24
	Total	166,1042	87,98797	48

Interpretasi

Dari Output 1 kita memperoleh rata-rata dan standar deviasi untuk menggambarkan tingkat gula darah dalam kelompok yang berbeda. Output 2 memberikan hasil analisis univariat ANOVA dua arah (di sini variabel yang diuji adalah kadar gula darah). Untuk menginterpretasikan hasil, kita perlu memahami hasil yang diberikan dalam Output 2. Judul kolom memberikan sumber varians, jumlah kuadrat, derajat kebebasan, rata-rata jumlah kuadrat, F (F-rasio) dan signifikansi. Dari sini, kita dapat menyatakan efek pengobatan. Kita melihat Value-F dan Signifikansi yang diberikan terhadap Eksperimen, Value-F adalah 169,784 dan signifikansi (p-value) adalah 0,000. Karena p-value kurang dari 0,01, maka ada perbedaan signifikan antara kelompok perlakuan.

Selanjutnya, kita dapat menyatakan apakah ada perbedaan yang signifikan antara laki-laki dan perempuan di mana kita mengambil Value yang diberikan terhadap Gender, di sini, Value F adalah 0,175 dan signifikansi adalah 0,678. Karena Value p lebih besar dari 0,01, perbedaan kadar gula darah antara pria dan wanita tidak signifikan.

Kemudian pengaruh interaktif dari kondisi eksperimental dan jenis kelamin pada tingkat gula darah ditafsirkan dari nilai-nilai yang diberikan terhadap Experimental*Gender, di mana rasio-F adalah 0,350 dan signifikansi adalah 0,790. Di sini p-value lebih besar dari 0,01; oleh karena itu, pengaruh interaktif Eksperimen dan Jenis Kelamin tidak signifikan. Karena itu, pria dan wanita bisa diperlakukan sama untuk menurunkan kadar gula darah.

Output 2

Tests of Between-Subjects Effects

Dependent Variable: Blood sugar level in mg/100ml

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	146154.312 ^a	7	20879.187	72.939	.000
Intercept	1324348.621	1	1324348.621	4628.478	.000
Experiment	145904.063	3	48601.354	169.764	.000
Gender	50.021	1	50.021	.175	.678
Experiment * Gender	300.229	3	100.076	.360	.790
Error	11450.167	40	286.254		
Total	1481953.000	48			
Corrected Total	157604.479	47			

a. R Squared = .927 (Adjusted R Squared = .915)

Output 3

3. Experimental groups * Gender

Dependent Variable: Blood sugar level in mg/100ml

Experimental groups	Gender	Mean	Std. Error	95% Confidence Interval	
Normal	Male	102.667	8.907	86.707	118.627
	Female	104.833	8.907	90.873	118.793
Diabetic	Male	243.333	8.907	229.373	257.293
	Female	253.500	8.907	239.540	267.460
Herb 1	Male	182.500	8.907	168.540	196.460
	Female	180.500	8.907	166.540	194.460
Herb 2	Male	131.833	8.907	117.873	145.793
	Female	129.667	8.907	115.707	143.627

Interpretasi

Output 3 memberikan rata-rata kadar gula darah dalam kelompok eksperimen yang berbeda sehubungan dengan jenis kelamin yang berbeda.

Output 4 memberikan kisaran kadar gula darah dalam kelompok eksperimen yang berbeda dengan mengabaikan jenis kelamin.

Output 4

Blood sugar level in mg/100ml

Duncan^{a,b}

Experimental groups	N	Subset			
Normal	12	103.7500			
Herb 2	12		130.7500		
Herb 1	12			181.5000	
Diabetic	12				248.4167
Sig.		1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = 286.254.
a. Uses Harmonic Mean Sample Size = 12.000.
b. Alpha = .05.

ANALISIS VARIANS GANDA – MULTIPLE ANOVA (MANOVA)

MANOVA adalah teknik yang menentukan efek variabel kategori independen pada beberapa variabel dependen. Ini biasanya digunakan untuk membandingkan beberapa kelompok variabel ganda

kontinyu atau kategori. Perbedaan utama antara MANOVA dan ANOVA adalah bahwa beberapa variabel dependen dipertimbangkan dalam MANOVA, sedangkan Uji ANOVA untuk perbedaan antar kelompok antara nilai rata-rata dari satu variabel dependen. MANOVA menggunakan satu atau lebih independen kategori sebagai prediktor, seperti ANOVA, tetapi tidak seperti ANOVA, terdapat lebih dari satu variabel dependen.

ANALISIS VARIANS GANDA - MULTIPLE ANOVA (MANOVA) DENGAN SPSS

Analisis varians multivariat (MANOVA) dalam SPSS mirip dengan ANOVA, sebagai pengganti satu variabel dependen metrik, kita memiliki dua atau lebih variabel dependen. MANOVA di SPSS berkaitan dengan perhitungan perbedaan antara kelompok dan perbedaan kelompok di beberapa variabel dependen secara bersamaan.

Misalkan empat kelompok, masing-masing terdiri dari 100 individu yang dipilih secara acak, terpapar dengan empat iklan komersial tentang beberapa deterjen. Setelah menonton iklan, setiap individu memberikan peringkat pada preferensi produk, untuk perusahaan manufaktur dan iklan komersial dicatat. Karena ketiga variabel ini berkorelasi, MANOVA dalam SPSS harus dilakukan untuk menentukan perusahaan yang menerima preferensi tertinggi di ketiga variabel (produk, perusahaan manufaktur dan iklan komersial). Mari kita memahami dasar-dasar melakukan MANOVA di SPSS, sebelum melanjutkan dengan proses yang sebenarnya melakukan hal yang sama.

DASAR-DASAR MELAKUKAN MANOVA DI SPSS

MANOVA di SPSS dilakukan dengan memilih **Analyze, General Linear Model** dan **Multivariate** dari menu utama. Seperti dalam ANOVA, langkah pertama adalah mengidentifikasi variabel dependen dan independen. MANOVA di SPSS melibatkan dua atau lebih variabel dependen metrik. Variabel metrik adalah variabel yang diukur menggunakan skala interval atau rasio. Variabel dependen umumnya dilambangkan oleh Y dan variabel independen dilambangkan dengan X.

Di sini, Hipotesis nol adalah rata-rata variabel dependen ganda yang sama di seluruh kelompok.

Seperti pada ANOVA, MANOVA dalam SPSS juga melibatkan dekomposisi total variasi yang diamati dalam semua variabel

dependen secara bersamaan. Variasi total dalam Y dilambangkan oleh SSY , yang dapat dipecah menjadi dua komponen:

$$SSY = SS_{\text{between}} + SS_{\text{within}}$$

Di sini subskrip 'between' dan 'within' mengacu pada kategori X. SS_{between} adalah bagian dari jumlah kuadrat di Y yang terkait dengan variabel independen atau faktor X. Jadi, umumnya disebut sebagai jumlah dari kotak X. SS_{within} adalah variasi dalam Y yang terkait dengan variasi dalam setiap kategori X. Hal ini umumnya disebut sebagai jumlah kuadrat untuk kesalahan dalam MANOVA. Jadi di MANOVA, untuk semua variabel dependen, misalnya, Y_1 , Y_2 (dan seterusnya), dekomposisi total variasi dilakukan secara bersamaan.

Tugas berikutnya di MANOVA dalam SPSS adalah untuk mengukur efek X pada Y_1 , Y_2 (dan seterusnya). Ini umumnya dilakukan dengan jumlah kuadrat X. Besaran relatif dari jumlah kuadrat X di MANOVA meningkat karena perbedaan antara rata-rata Y_1 , Y_2 (dan seterusnya) dalam kategori X meningkat. Besaran relatif dari jumlah kuadrat X di MANOVA menurun karena variasi dalam Y_1 , Y_2 (dan seterusnya) dalam kategori X menurun.

Langkah terakhir dalam MANOVA dalam SPSS adalah menghitung kuadrat rata-rata yang diperoleh dengan membagi jumlah kuadrat dengan derajat kebebasan yang sesuai. Hipotesis nol rata-rata yang sama diuji oleh statistik F, yang merupakan rasio dari kuadrat rata-rata terkait dengan variabel independen (MS_{between}) ke kuadrat rata-rata terkait dengan kesalahan (MS_{within}).

Contoh 5.3

Lakukan analisis varians multivariat (MANOVA) pada data konsentrasi plasma kalsium (dalam mg/100 ml) dan untuk tingkat kehilangan air evaporatif (dalam mg/menit).

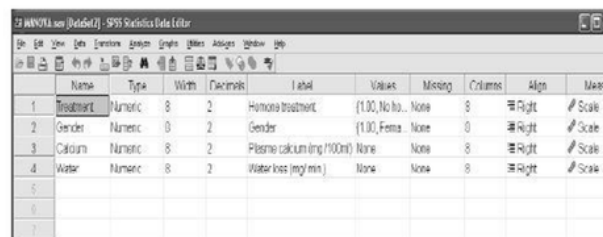
No hormone treatment				Hormone treatment			
Female		Male		Female		Male	
Plasma calcium	Water loss	Plasma calcium	Water loss	Plasma calcium	Water loss	Plasma calcium	Water loss
16.5	76	14.5	80	39.1	71	32.0	65
18.4	71	11.0	72	26.2	70	23.8	69
12.7	64	10.8	77	21.3	63	28.8	97
14.0	66	14.3	69	35.8	59	25.0	56
12.8	69	10.0	74	40.2	60	29.3	52

Langkah 1 Buka Data Editor Data, klik Variabel view dan ikuti Langkah 1-4 untuk memberi nama variabel (Gambar 5.12). Di kolom pertama, ketik "Treatment", pilih Numerik pada Type, beri label sebagai "Hormone treatment" di Label. Di kolom Value klik pada area abu-abu, jendela popup terbuka (detail dapat dimasukkan seperti pada contoh sebelumnya), ketik "1" di kotak Value dan "No hormone treatment" di Label dan klik Add untuk mentransfernya ke kotak bawah. Ketik "2" di kotak Value dan "Hormon treatment" di Label dan klik Add untuk mentransfernya ke kotak bawah dan klik OK.

Langkah 2 Di baris kedua, ketik "Gender" pada Name, pilih Numerik pada Type, beri label "Gender" di Label. Di kolom Value, klik pada area abu-abu, di jendela popup ketik 1 di kotak Value dan "Female" di kotak Label dan klik Add untuk mentransfernya ke kotak bawah. Ketik "2" di kotak Value dan "Male" di kotak Label, klik Add dan klik OK.

Langkah 3 Pada baris ketiga, ketik "Calcium" di bawah Nama, pilih Numeric pada Type, beri label "Plasma calcium (mg/100 ml)" di Label.

Langkah 4 Di baris keempat, ketik "Water" di bawah Name, pilih Numeric pada Type, beri label "Water loss (mg/100min)" pada Label.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Treatment	Numeric	8	2	Hormone treatment	(1,0, No ho... None		8	Right	Scale
2	Gender	Numeric	8	2	Gender	(1,0, Feme... None		8	Right	Scale
3	Calcium	Numeric	8	2	Plasma calcium (mg/100ml)	None	None	8	Right	Scale
4	Water	Numeric	8	2	Water loss (mg/min)	None	None	8	Right	Scale

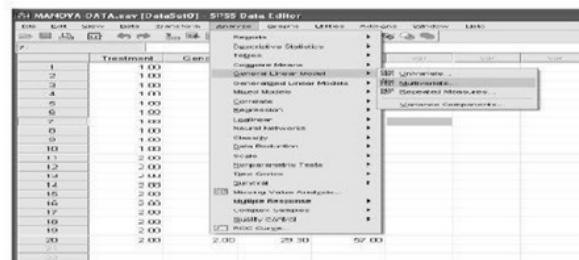
Gambar 5.13 Variabel view dengan rincian yang dimasukkan untuk empat variabel

Langkah 5 Klik Data view dan ketik data dalam Data view di setiap judul (Gambar 6.14). Hati-hati untuk memasukkan data sedemikian rupa sehingga dua kategori "1" dimasukkan untuk No hormone treatment dan "2" dimasukkan untuk hormon treatment di kolom pertama. Di kolom kedua masukkan "1" untuk female dan "2" untuk male terhadap setiap perlakuan dengan benar. Di kolom ketiga masukkan level calsium sesuai dengan hormon treatment dan gender. Di kolom keempat masukkan water loss sesuai dengan hormon treatment dan gender.

	Treatment	Gender	Creatinine	Water loss
1	1.00	1.00	10.50	75.00
2	1.00	1.00	10.40	71.00
3	1.00	1.00	12.70	84.00
4	1.00	1.00	14.00	65.00
5	1.00	1.00	12.80	69.00
6	1.00	2.00	14.50	80.00
7	1.00	2.00	11.00	73.00
8	1.00	2.00	10.90	77.00
9	1.00	2.00	14.30	69.00
10	1.00	2.00	10.00	70.00
11	2.00	1.00	20.10	74.00
12	2.00	1.00	20.20	70.00
13	2.00	1.00	21.30	63.00
14	2.00	1.00	35.00	59.00
15	2.00	1.00	40.20	60.00
16	2.00	2.00	32.00	65.00
17	2.00	2.00	23.00	60.00
18	2.00	2.00	25.00	97.00
19	2.00	2.00	25.00	55.00
20	2.00	2.00	29.30	52.00

Gambar 5.14 Data view dengan data yang dimasukkan

Langkah 6 Dari menu utama pilih Analyze, klik General Linear Model dan kemudian pilih Multivariate (Gambar 5.15).



Gambar 5.15 Memilih Multivariate dari menu drop-down

Langkah 7 Jendela popup dengan judul Multivariate terbuka seperti pada Gambar 6.16.



Gambar 5.16 Kotak dialog multivariate untuk mentransfer variabel

Langkah 8 Pindahkan variabel "Plasma calsium mg/100ml" dan "Water loss mg/min" untuk Variabel Dependens: dan Treatment dan Gender untuk Fixed Factor:(Gambar 5.17). Karena kita hanya memiliki dua kategori pada Treatment dan Gender, kita tidak dapat melakukan pengujian post-hoc untuk kumpulan data ini.



Gambar 5.17 Multivariate: dengan variabel dependen dan fixed factor ditransfer

Langkah 9 Klik Opsi untuk membuka kotak dialog Multivariate: Opsi. Transfer (OVERALL) dari Factor(s) dan Faktor Interactions: ke Display Means for. Pilih kotak centang Statistik Deskriptif pada Display (Gambar 5.18). Klik Lanjutkan untuk kembali ke kotak Multivariate.



Gambar 5.18 Multivariate: Options untuk memilih Statistik Deskriptif

Langkah 10 Klik OK untuk menjalankan analisis dan mendapatkan output.

Output muncul di bawah judul yang berbeda, yaitu: **Between Subject Factors, Descriptive Statistics, Multivariate Tests and Tests of Between-Subjects Effects**. Namun kita menggunakan Statistik Deskriptif (Output 1) dan Tests of Between-Subjects Effects (Output 2) untuk contoh ini.

Descriptive Statistics					
	Treatment	Gender	Mean	Std. Deviation	N
Plasma Calcium in mg /100ml	No hormone	Female	14.8800	2.49339	5
		Male	12.1200	2.11600	5
		Total	13.5000	2.62066	10
	Hormone	Female	28.8600	5.52883	5
		Male	28.3800	4.39454	5
		Total	28.6200	4.71618	10
	Total	Female	21.6700	8.46463	10
		Male	20.2500	9.16591	10
		Total	21.0600	8.55923	20
Water loss in mg/min	No hormone	Female	69.2000	4.66633	5
		Male	74.4000	4.27765	5
		Total	71.8000	5.02881	10
	Hormone	Female	64.6000	5.59484	5
		Male	62.8000	5.93296	5
		Total	63.7000	5.51666	10
	Total	Female	65.9000	5.42525	10
		Male	69.6000	7.87020	10
		Total	67.7600	6.66641	20

Interpretasi

Dari Output 1, Statistik Deskriptif kita dapat menggambarkan tingkat kalsium darah dan kehilangan air dalam kelompok eksperimen yang berbeda dari laki-laki dan perempuan. Kita dapat menyimpulkan bahwa tingkat kalsium darah pada wanita yang tidak menerima hormon adalah $14,88 \pm 2,49$ mg/100ml dan laki-laki adalah $12,12 \pm 2,11$ mg/100ml. Tingkat kalsium darah pada perempuan yang menerima hormon adalah $28,86 \pm 5,52$ mg/100 ml dan laki-laki adalah $28,38 \pm 4,39$ mg/100ml. Demikian pula, kehilangan air pada wanita yang tidak menerima hormon adalah $69,20 \pm 4,65$ mg/menit dan pada pria adalah $74,40 \pm 4,28$ mg/menit. Kehilangan air pada wanita yang menerima hormon adalah $64,6 \pm 5,59$ mg/menit dan pada wanita adalah $62,8 \pm 5,93$ mg/menit.

Output 2

Tests of Between-Subjects Effects						
Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Plasma Calcium in mg /100ml	1182.632 ^a	3	394.544	75.593	.000
	Water loss in mg/min	439.739 ^a	3	146.579	5.055	.012
Intercept	Plasma Calcium in mg /100ml	6070.472	1	6070.472	585.781	.000
	Water loss in mg/min	91801.250	1	91801.250	3.446E7	.000
Treatment	Plasma Calcium in mg /100ml	1144.672	1	1144.672	75.489	.000
	Water loss in mg/min	328.050	1	328.050	12.321	.003
Gender	Plasma Calcium in mg /100ml	19.422	1	19.422	.697	.360
	Water loss in mg/min	14.450	1	14.450	.543	.472
Treatment * Gender	Plasma Calcium in mg /100ml	8.408	1	8.408	.429	.522
	Water loss in mg/min	61.250	1	61.250	2.300	.140
Error	Plasma Calcium in mg /100ml	242.256	16	15.143		
	Water loss in mg/min	436.000	16	27.250		
Total	Plasma Calcium in mg /100ml	10276.460	20			
	Water loss in mg/min	92631.000	20			
Corrected Total	Plasma Calcium in mg /100ml	1404.888	19			
	Water loss in mg/min	618.750	19			

a. R Squared = .826 (Adjusted R Squared = .799)

b. R Squared = .467 (Adjusted R Squared = .393)

Dari hasil **2, Tests of Between-Subjects Effects**, kita dapat menyimpulkan pengaruh terapi hormon, jenis kelamin dan pengaruh interaktif dari keduanya (hormon dan jenis kelamin) pada kalsium darah dan kehilangan air. Ada 7 kolom dan 8 baris dalam output. Kolom pertama memberi sumber variasi. Dari kolom itu kita dapat

memformulasikan tabel ANOVA dengan mengambil nilai-nilai terhadap Pengobatan, Jenis Kelamin dan Pengobatan*Jenis Kelamin, Kesalahan dan Total. Kita dapat mengambil hasil yang diberikan terhadap Perawatan, Gender dan Pengobatan * Jenis Kelamin untuk menjawab masalah kita.

F value untuk pengobatan kalsium plasma adalah 75,48 dan signifikansinya adalah 0,000 ($p < 0,01$). F value untuk dan untuk kehilangan air adalah 12,32 dan signifikansi 0,003 ($p < 0,05$).

Oleh karena itu, terapi hormon memiliki efek yang signifikan pada kadar kalsium plasma dan kehilangan air.

F value untuk gender untuk kalsium plasma adalah 0,867 dan signifikansi 0,336 ($p > 0,05$). F value untuk kehilangan air adalah 0,543 dan signifikansi adalah 0,472 ($p > 0,05$). Oleh karena itu, kadar kalsium plasma dan kehilangan air tidak berbeda secara signifikan pada pria dan wanita atau jenis kelamin tidak memiliki pengaruh pada tingkat kalsium dan kehilangan air.

F value untuk pengaruh interaktif pengobatan dan jenis kelamin pada kalsium plasma adalah 0,429 dan signifikansi adalah 0,522 ($p > 0,05$). F value untuk kehilangan air adalah 2,3 dan signifikansi adalah 0,149 ($p > 0,05$). Oleh karena itu, pengaruh interaktif dari kedua hormon dan jenis kelamin tidak signifikan. Hasilnya dapat disimpulkan dengan mengatakan bahwa pria dan wanita tidak berbeda secara signifikan dalam kadar kalsium plasma mereka dan kehilangan air dan perawatan hormon hanya memiliki pengaruh yang signifikan pada keduanya. Oleh karena itu, pria dan wanita dapat diperlakukan sama dengan hormon untuk mengontrol kalsium plasma dan kehilangan air.

REVIEW LATIHAN

1. Dua sampel diambil dari dua populasi normal dan data berikut diperoleh. Ujilah apakah kedua sampel memiliki varian yang sama pada tingkat signifikansi 5%.

Sample 1	60	65	71	74	76	82	85	87
Sample 2	61	66	67	85	78	63	86	85

2. Kandungan Klorida (mg/lit) dari sungai di situs tertentu yang dikumpulkan dari empat musim yang berbeda diberikan di bawah ini. Hitunglah rata-rata hitung dan standar deviasi. Analisis signifikansi varians. Tentukan keragaman kandungan oksigen untuk empat musim yang berbeda. Hitung rentang dalam konten klorida untuk musim yang berbeda.

Winter	Summer	Monsoon	Post monsoon
68	102	99	88
69	103	98	90
70	110	100	83
71	106	86	91
72	105	88	80
75	108	89	82

3. Di sungai di mana limbah diizinkan untuk dicampur, nilai BOD (oksigen dalam mg/L) dihitung di lokasi yang berbeda dan diberikan di bawah ini. Hitung rata-rata hitung dan standar deviasi. Lakukan analisis signifikansi varians antara situs yang berbeda.

Site A	Site B	Site C	Site D
1.2	2.2	3.2	4.3
1.0	2.3	2.9	3.9
1.3	2.2	2.1	3.6
1.0	2.3	3.0	3.4
1.0	2.0	2.9	3.5
1.1	1.9	2.8	3.9
1.2	1.8	3.3	3.8

4. Tingkat hemoglobin dari empat kelompok anak yang diberi empat diet yang berbeda diberikan di bawah ini. Hitung rata-rata hitung dan standar deviasi. Lakukan analisis signifikansi varian dan rentang dalam respon diet pada hemoglobin.

Diet A	Diet B	Diet C	Diet D
11.6	13.2	10.1	12.2
10.6	13.1	10.1	13.1
11.1	13.	9.9	13.
11.2	14.2	9.8	12.2
10.3	13.5	10.1	12.5
11.5	12.8	10.1	11.8
12.1	13.0	11.1	13.0
12.6	12.5	10.5	12.5

5. Data berikut ini memberikan produksi gandum dalam ton/hektar dari tiga varietas yang berbeda A, B dan C. Apakah ada perbedaan yang signifikan dalam produksi tiga varietas?

Variety A	Variety B	Variety C
68	102	99
69	103	98
70	110	86
71	106	86
72	105	88
75	108	89

Bab 6

KORELASI

ASOSIASI STATISTIK ANTARA VARIABEL

Dalam situasi kehidupan nyata, hubungan atau asosiasi terjadi antara karakter atau variabel. Sebagai contoh, jika kita mengukur tinggi dan berat dari sekelompok individu, kita memperoleh suatu seri di mana setiap individu dari seri memiliki 2 nilai, satu berkaitan dengan tinggi dan yang lain berhubungan dengan berat. Distribusi seperti itu di mana setiap individu (unit) memiliki dua nilai disebut distribusi bivariat. Jika kita mengukur lebih dari dua variabel pada setiap unit distribusi disebut distribusi multivariat. Dalam keduanya, perubahan dalam satu variabel ditemukan atau ternyata ditemukan terkait dengan perubahan dalam variabel lain. Hubungan ini mungkin kasual atau kausal. Statistik sebagai alat yang digunakan dalam hubungan tersebut masing-masing adalah korelasi dan regresi. Ada banyak jenis data dalam ilmu biologi dan ilmu sosial di mana hubungan antara dua variabel bukanlah salah satu ketergantungan. Dalam kasus seperti itu, besarnya salah satu variabel berubah ketika besarnya perubahan variabel kedua, tetapi tidak masuk akal untuk menganggapnya sebagai variabel independen atau dependen. Dalam situasi seperti itu, analisis korelasi diperlukan.

Contoh data yang cocok untuk analisis korelasi adalah pengukuran panjang lengan dan panjang kaki manusia. Dapat ditemukan bahwa seorang individu dengan lengan panjang biasanya memiliki kaki yang panjang, sehingga hubungan mungkin diinginkan; tetapi tidak ada pembenaran dalam menyatakan bahwa panjang satu anggota bergantung pada panjang yang lain. Tetapi untuk variabel seperti usia dan tekanan darah orang, usia mempengaruhi tekanan darah. Analisis regresi adalah yang paling cocok untuk jenis variabel ini, meskipun kita dapat menerapkan analisis korelasi untuk menemukan tingkat asosiasi. Sejauh mana dua atau lebih variabel co-variasi dalam beberapa BENTUK linier diberikan oleh analisis korelasi.

21 Jadi dua variabel dikatakan berkorelasi jika perubahan dalam satu variabel menghasilkan perubahan terkait dalam variabel lain. Oleh karena itu, korelasi adalah alat statistik yang mempelajari hubungan antara 2 variabel.

KORELASI SEDERHANA DAN KORELASI BERGANDA

Dua variabel, misalnya, usia dan tekanan darah ditemukan bervariasi dalam beberapa BENTUK linier. Cara yang tepat untuk menyatakan tingkat hubungan linear antara dua variabel ini adalah dengan menghitung koefisien korelasi. Oleh karena itu, ketika seseorang mempelajari hubungan antara dua variabel itu disebut analisis korelasi sederhana dan jika lebih dari dua variabel yang terlibat adalah analisis multivariat.

Jenis Korelasi

Korelasi positif dan negatif Jika nilai dari dua variabel bergerak dalam arah yang sama, yaitu, jika peningkatan nilai dari satu variabel menghasilkan peningkatan yang sesuai dalam nilai-nilai variabel lain atau jika penurunan nilai dari satu variabel menghasilkan penurunan yang sesuai dalam nilai-nilai dari variabel lain, ada korelasi positif.

Di sisi lain, jika variabel menyimpang dalam arah yang berlawanan, yaitu, jika peningkatan nilai dari satu variabel menghasilkan penurunan yang sesuai dalam Nilai-nilai variabel lain, terdapat korelasi negatif.

Korelasi linier dan non-linear Jika perubahan unit dalam satu variabel berhubungan dengan perubahan konstan dalam variabel lain di seluruh rentang data, maka korelasi antara dua variabel dikatakan linier. Sebagai contoh, data berikut menunjukkan korelasi linear antara dua variabel X dan Y.

X	1	2	3	4	5
Y	5	7	9	11	13

Jadi untuk perubahan unit dalam Nilai X, ada perubahan konstan yaitu, 2 dalam Nilai-nilai yang sesuai dari Y. Ketika data ini diplot pada lembar grafik, ini memberikan garis lurus. Tetapi kondisi ideal seperti itu tidak pernah ada di alam. Dalam kebanyakan situasi dalam ilmu sosial dan biologi hubungan antara 2 variabel tidak linier dan dapat berfluktuasi.

Hubungan antara dua variabel dikatakan non-linear atau curvilinear, jika perubahan unit dalam satu variabel tidak

sesuai dengan perubahan dalam variabel lain pada tingkat yang konstan tetapi pada tingkat yang berfluktuasi. Jenis data seperti itu, ketika diplot pada lembar grafik, kita tidak mendapatkan garis lurus. Secara matematis, korelasi dikatakan non-linear jika kecenderungan kurva diplot tidak konstan.

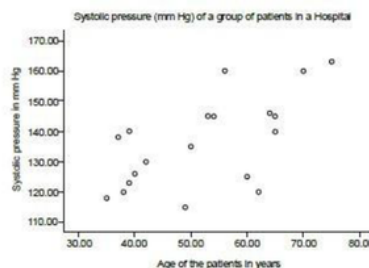
Teknik untuk analisis dan pengukuran hubungan non-linear cukup berbelit-belit dan rumit jika dibandingkan dengan metode belajar dan mengukur hubungan yang linier.

METODE STUDI KORELASI

Ada berbagai metode untuk mempelajari tingkat korelasi yang tergantung pada jenis variabel dan jumlah variabel yang terlibat. Jika hanya dua variabel yang dipelajari dalam perhitungan, digunakan analisis bivariat. Tingkat hubungan antara dua variabel pengukuran dapat dilakukan secara grafis dan matematis.

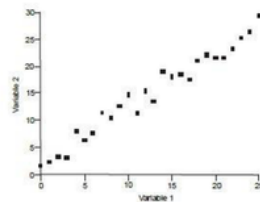
METODE GRAFIK: DIAGRAM PENCAR

Cara paling sederhana untuk memastikan korelasi antara dua variabel adalah dengan diagram pencar. Dalam metode ini, jika n pasang Nilai ($X_1Y_1, X_2Y_2, \dots, X_nY_n$) dari dua variabel X dan Y diberikan, maka satu diwakili sepanjang absis (sumbu X) dan yang lainnya sepanjang koordinat (sumbu Y) pada lembar grafik. Dalam contoh dari dua variabel, yaitu usia orang dan tekanan darah, untuk setiap kasus atau setiap orang satu titik diplot untuk usia dan tekanan darahnya (Umur diambil di sumbu- X dan tekanan darah di Sumbu- Y). Ketika semua pasangan " n " diplot untuk seluruh rangkaian data, diagram titik-titik yang diperoleh disebut diagram pencar (Scatter). Diagram pencar untuk data yang sama diberikan di bawah ini. Dari diagram pencar tersebut kita dapat membentuk gagasan yang cukup bagus tetapi kasar tentang hubungan antara dua variabel.

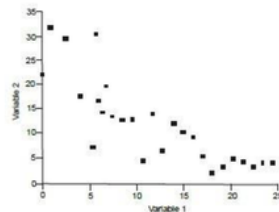


Gambar 6.1 Diagram pencar tekanan sistolik (mm Hg) dan usia (tahun)

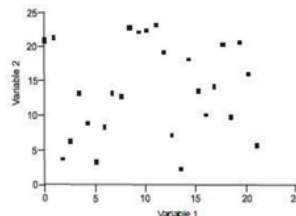
Dalam diagram pencar jika titik mulai dari sudut kiri bawah dan memanjang ke sudut kanan atas seperti yang ditunjukkan pada Gambar 6.2, ada korelasi positif antara dua variabel ini. Di sisi lain, ketika poin bergerak dari sudut kanan bawah ke pojok kiri atas seperti yang ditunjukkan pada Gambar 6.3, ada korelasi negatif antara dua variabel ini. Jika poin tersebar di seluruh seperti pada Gambar 6.4, maka tidak ada korelasi.



Gambar 6.2 Diagram pencar yang menunjukkan korelasi positif antar variable



13
Gambar 6.3. Diagram pencar yang menunjukkan korelasi negatif antara dua variabel



Gambar 6.4 Diagram pencar tidak ada korelasi antara dua variabel

METODE MATEMATIKA: KOEFISIEN KORELASI PEARSON

Hubungan antara dua variabel yang pengukurannya dipelajari dengan metode yang diperkenalkan oleh Karl Pearson dan disebut Koefisien Korelasi Pearson. Rumus berikut digunakan untuk menemukan hubungan antara dua variabel:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

dimana,

r – koefisien korelasi

x_i – Nilai item ke- i dari variabel x

\bar{x} rata-rata variabel x

y_i - nilai ke- i item variable y

\bar{y} rata-rata variabel y and Σ -sum

Nilai koefisien korelasi Pearson jatuh antara +1 dan -1. Jika variabel berkorelasi negatif nilai r terletak antara 0 dan -1, jika mereka berkorelasi positif nilainya adalah antara 0 dan +1. Jika tidak ada korelasi, nilainya adalah 0. Tetapi nilai dari 0 ke 1 ditafsirkan berdasarkan Uji signifikansi.

ANALISIS BIVARIAT DENGAN SPSS

Perhitungan koefisien korelasi dengan rumus matematika menjadi membosankan ketika seseorang suka menghitung lebih banyak jumlah variabel dari sampel yang sangat besar. SPSS memungkinkan kita melakukan analisis untuk sejumlah besar variabel sekaligus. Mari kita cari koefisien korelasi untuk sepasang variabel pertama dengan contoh yang sudah disebut di atas.

Contoh 6.1

Carilah koefisien korelasi untuk variabel, usia (tahun) dan tekanan darah sistolik (mmHg) pada manusia.

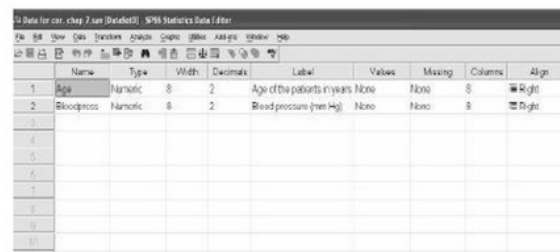
Age	56	42	60	50	54	49	39	62	65	70	40	53	35	38	39	37	70	75	65	64
BP	160	130	125	135	145	115	140	120	140	160	126	145	118	120	123	138	160	163	145	146

Sebelum masuk ke SPSS, penting untuk membuat Hipotesis nol dan Hipotesis alternatif.

Hipotesis nol (H_0): Tidak ada korelasi antara usia dan tekanan darah sistolik.

Hipotesis alternatif (H_A) Ada korelasi antara usia dan tekanan darah sistolik.

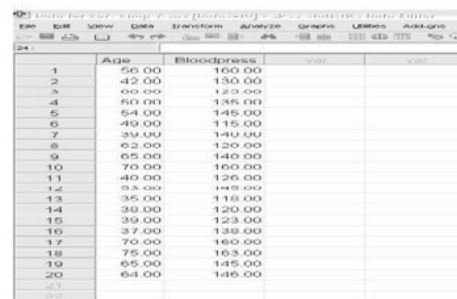
Langkah 1 Buka Data Editor dan klik Variabel view lalu masukkan nama variabel dan detail variabel (Gambar 7.5).



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	Age	Numeric	8	2	Age of the patients in years	None	None	8	Right
2	Bloodpress	Numeric	8	2	Blood pressure (mm Hg)	None	None	8	Right

Gambar 6.5 Variabel view dengan dua variabel yang dimasukkan

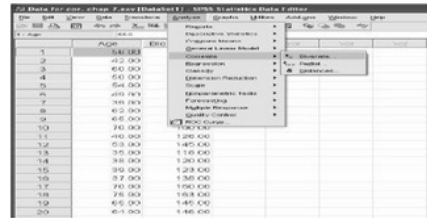
Langkah 2 Klik Data view dan masukkan data pada variabel yang sesuai (Gambar 6.6).



	Age	Bloodpress
1	56.00	160.00
2	42.00	130.00
3	60.00	125.00
4	50.00	135.00
5	54.00	145.00
6	49.00	115.00
7	39.00	140.00
8	62.00	120.00
9	65.00	140.00
10	70.00	160.00
11	40.00	126.00
12	53.00	145.00
13	35.00	118.00
14	38.00	120.00
15	39.00	123.00
16	37.00	138.00
17	70.00	160.00
18	75.00	163.00
19	65.00	145.00
20	64.00	146.00

Gambar 6.6 Tampilan Data dengan data untuk dua variabel

Langkah 3 Pilih Analisis dari menu utama dan pilih **Correlate** dan kemudian pilih **Bivariate** seperti yang ditunjukkan pada Gambar 6.7.



Gambar 6.7 Memilih opsi Bivariat dari menu utama

Langkah 4 kotak dialog Korelasi Bivariat terbuka seperti pada Gambar 6.8.



Gambar 6.8 Dialog Korelasi Bivariat

Langkah 5 Transfer variabel ke dalam kotak Variabel, pilih Pearson di bawah Koefisien Korelasi dan pilih **Two-tailed** pada **Significance and Flag Significant Correlations** (Gambar 6.9).



Gambar 6.9 Korelasi Bivariat dengan opsi yang dipilih untuk koefisien korelasi Pearson

Langkah 6 Klik Opsi untuk membuka Korelasi Bivariat: Opsi, pilih Mean dan Standar Deviasi di bawah Statistik, sehingga Anda bisa

mendapatkan beberapa statistik deskriptif dalam output, meskipun opsi ini tidak diperlukan untuk menjalankan koefisien korelasi Pearson yang sebenarnya (Gambar 6.10).



Gambar 6.10 Korelasi Bivariat: untuk memilih rata-rata hitung dan standar deviasi

Langkah 7 Klik Lanjutkan dan klik OK untuk menjalankan analisis. Dua output, satu untuk statistik deskriptif (Output 1) dan yang lainnya untuk Korelasi (Output 2) muncul seperti ditunjukkan di bawah ini:

Output 1

Descriptive Statistics			
	Mean	Std. Deviation	N
Age of the patients in years	53.1500	12.81970	20
Blood pressure (mm Hg)	137.7000	15.35921	20

Output 2

Correlations			
		Age of the patients in years	Blood pressure (mm Hg)
Age of the patients in years	Pearson Correlation	1	.661**
	Sig. (2-tailed)		.002
	N	20	20
Blood pressure (mm Hg)	Pearson Correlation	.661	1
	Sig. (2-tailed)	.002	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Interpretasi

Output 1 memberikan usia rata-rata dan tekanan darah dengan standar deviasi.

Output 2 memberikan matriks utama dari koefisien korelasi Pearson. Variabel telah diatur dalam suatu matriks sedemikian rupa sehingga kolom/baris mereka berpotongan. Di dalam sel ada angka yang menceritakan tentang interaksi statistik antar variabel. Tiga jenis informasi disediakan di setiap sel, yaitu, Korelasi Pearson, signifikansi dan jumlah kasus. Nilai-nilai di kedua sisi diagonal adalah menjelaskan satu sama lain, yaitu, nilainya sama. Oleh karena itu, orang dapat mengabaikan informasi di atas diagonal atau di bawah diagonal jika diinginkan.

Nilai Korelasi Pearson, yaitu 0.661 adalah nilai-r. Karena, nilai-r adalah nilai positif dan signifikan (2-tailed) sebesar 0,002 (p-Value) di bawah 0,01, kita menolak H_0 dan menyimpulkan bahwa peningkatan usia benar-benar meningkatkan tekanan darah sistolik atau usia pasien dan tekanan darah berhubungan positif.

Contoh 6.2

Berikut ini adalah data yang terkait dengan nilai ujian yang diperoleh oleh mahasiswa dengan waktu belajar mereka dan ketidakhadiran mereka di kelas. Lakukan analisis bivariat untuk data berikut:

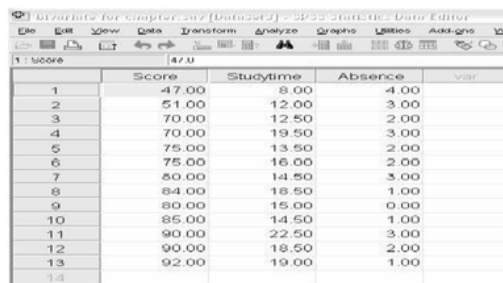
Case no.	Test score	Study time in hours/week	Absence in hours/week
1	47	8	4
2	51	12	3
3	70	12.5	2
4	70	19.5	3
5	75	13.5	2
6	75	16	2
7	80	14.5	3
8	84	18.5	1
9	80	15	0
10	85	14.5	1
11	90	22.5	3
12	90	18.5	2
13	92	19	1

Dalam latihan ini ada tiga variabel, kita dapat menghitung korelasi antara nilai ujian dan waktu belajar serta nilai ujian dan ketidakhadiran di kelas. SPSS memungkinkan kita untuk melakukan analisis bivariat karena banyak variabel dimana kita ingin menemukan hubungannya. Dalam situasi seperti ini kita dapat mengajukan Hipotesis nol untuk dua variabel secara terpisah. Di sini, kita mengajukan Hipotesis hanya untuk dua variabel pada suatu waktu, seperti nilai ujian dan waktu belajar. Demikian pula kita dapat mengusulkan untuk Nilai Ujian

dan ketidakhadiran di kelas dan waktu belajar dan ketidakhadiran di kelas. Nol dan Hipotesis alternatif diusulkan secara individual untuk setiap pasangan.

Langkah 1 Buka Data Editor, dan Klik Variabel View. Masukkan nama tiga variabel dan detail untuk masing-masing variabel.

Langkah 2 Klik Data View dan masukkan data di bawah variabel yang sesuai (Gambar 6.11).



	Score	Studytime	Absence	var
1	47.00	8.00	4.00	
2	51.00	12.00	3.00	
3	70.00	12.50	2.00	
4	70.00	19.50	3.00	
5	75.00	13.50	2.00	
6	75.00	16.00	2.00	
7	80.00	14.50	3.00	
8	84.00	18.50	1.00	
9	80.00	15.00	0.00	
10	85.00	14.50	1.00	
11	90.00	22.50	3.00	
12	90.00	18.50	2.00	
13	92.00	19.00	1.00	

Gambar 6.11 Tampilan Data dengan data untuk tiga variabel

Langkah 3 Pilih Analyze dari menu utama dan pilih **Correlate** dari menu dan **Bivariate**. Kotak dialog Korelasi Bivariat muncul.

Langkah 4 Sekarang pindahkan variabel dari sisi kiri ke kotak Variabel di sebelah kanan dengan mengklik panah.

Langkah 5 Pilih Pearson di bawah Koefisien Korelasi dan pilih **Two-tailed** pada Uji signifikansi.

Langkah 6 Klik Opsi dan pilih kotak Rata-rata dan Penyimpangan Standar di bawah Statistik.

Langkah 7 Klik Lanjutkan dan klik OK untuk menjalankan analisis.

Langkah 8 Output 1 dan 2 muncul seperti yang diberikan di bawah ini:

Output 1

	Mean	Std. Deviation	N
Test score obtained by students	76.0769	14.03841	13
Study time in hours / week	15.6923	3.85972	13
Absence in the class in hours / week	2.0769	1.11516	13

Output 2

Correlations				
		Test score obtained by students	Study time in hours / week	Absence in the class in hours / ...
Test score obtained by students	Pearson Correlation Sig. (2-tailed) N	1 .002 13	.777** .002 13	.607* .031 13
Study time in hours / week	Pearson Correlation Sig. (2-tailed) N	.777** .002 13	1 .002 13	-.255 .400 13
Absence in the class in hours / week	Pearson Correlation Sig. (2-tailed) N	-.597* .031 13	-.255 .400 13	1 .400 13

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Interpretasi

1. Output 1 memberikan rata-rata hitung, standar deviasi dan jumlah kasus. Dengan ini kita bisa mendeskripsikan data.

2. Output 2 adalah dalam bentuk matriks. Nilai-nilai di kedua sisi diagonal adalah menjelaskan satu sama lain, yaitu nilai-nilainya sama. Kita bisa merujuk dan menginterpretasikan nilai-nilai pada satu sisi diagonal. Nilai r positif untuk "variables test scores" (nilai ujian) dan "studying time" (waktu belajar), p -Value 0,003 di bawah 0,01, jadi kita menolak H_0 . Oleh karena itu, ada korelasi positif yang signifikan antara keduanya. Dengan demikian, peningkatan waktu belajar benar-benar meningkatkan Nilai Ujian.

3. Nilai r adalah -0,637 untuk variable Nilai Ujian dan ketidakhadiran di kelas dan p -Value (0,026) di bawah 0,01, jadi kita menolak H_0 , oleh karena itu ada korelasi yang signifikan antara keduanya, yaitu ketidakhadiran di kelas benar-benar berkorelasi negatif dengan Nilai Ujian.

4. Nilai r -0,360 negatif, untuk variabel waktu dan ketidakhadiran di kelas dan Nilai p adalah 0,25, yang lebih besar dari 0,01 sehingga kita tidak menolak H_0 , yaitu, tidak ada korelasi antara waktu belajar dan ketidakhadiran di kelas dan oleh karena itu, waktu belajar benar-benar tidak berkorelasi dengan ketidakhadiran di kelas dan atau sebaliknya.

KORELASI RANK (PERINGKAT)

Kadang-kadang, kita menemukan data statistik di mana variabel yang diteliti tidak mampu mengukur secara kuantitatif yang akurat dan mungkin tidak jatuh dalam skala interval juga, tetapi dapat diatur dalam urutan serial. Variabel tersebut adalah karakteristik atau atribut kualitatif seperti kecerdasan, kecantikan, moralitas, metode pengajaran, kejujuran, dll., yang tidak dapat diukur secara langsung dengan kuantitatif tetapi dapat disusun secara serial. Untuk data seperti itu, koefisien korelasi Karl Pearson tidak dapat digunakan. Salah satu metode menganalisis data semacam itu adalah dengan memberi peringkat variasi dan menghitung koefisien korelasi peringkat. Jadi korelasi peringkat adalah studi tentang hubungan antara peringkat yang berbeda pada set item yang sama. Ini berkaitan dengan hubungan antara dua peringkat dan menetapkan signifikansi hubungan ini. Data yang cocok untuk metode peringkat adalah yang tidak dapat diukur pada skala absolut, tetapi hanya pada skala ordinal. Ada metode yang berbeda dari data peringkat dan menghitung koefisien korelasi. Kita akan membahas tentang koefisien korelasi peringkat Spearman dan koefisien korelasi Kendall.

KOEFISIEN KORELASI RANK SPEARMAN'S

Charles Edward Spearman, seorang Psikolog Inggris mengembangkan formula untuk memperoleh koefisien korelasi antara jajaran sekelompok individu untuk sepasang atribut tertentu. Untuk menghitung koefisien korelasi Spearman, peringkat ditentukan untuk satu set variabel X dan Y secara individual. Dalam peringkat, pertama variabel X di peringkat dari rendah ke tinggi (bisa dari tinggi ke rendah juga); maka variabel Y diberi peringkat dengan cara yang sama. Penyimpangan antara jajaran pasangan variabel adalah, dikuadratkan dan dijumlahkan. Nilai-nilainya diganti dalam rumus berikut untuk mendapatkan koefisien korelasi Pearson. Karena deviasi antara peringkat-peringkat yang membentuk dasar perhitungan koefisien korelasi Pearson, biasanya disebut sebagai korelasi Rank.

$$R = 1 - \frac{6(\sum d^2)}{n(n^2 - 1)}$$

dimana,

R – Peringkat koefisien korelasi

d^2 – deviasi kuadrat

Σ –jumlah

n – pasangan item-item

Interpretasi

Nilai R terletak antara -1 dan +1 dan diinterpretasikan seperti dalam kasus koefisien korelasi Pearson.

KORELASI RANK SPEARMAN DENGAN SPSS

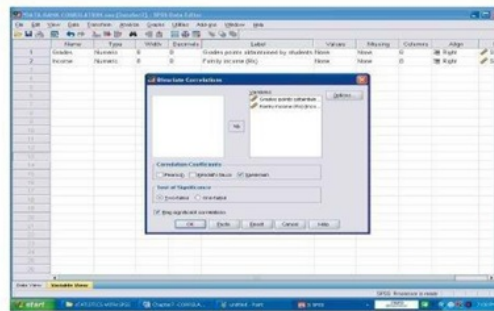
Contoh 6.3

Hitung koefisien korelasi peringkat Spearman untuk data yang diberikan di bawah ini untuk data pencapaian akademik dan pendapatan keluarga.

Grade points	75	73	96	61	71	56	50	85	90	54
Family income (Rs)	8500	7000	6000	12000	12000	5000	18000	9000	7000	8200

Langkah 1 Beri nama variabel dalam Variabel View dan data dalam Data editor

Langkah 2 Klik Analisis, lalu pilih Korelasi dan klik Bivariat. Korelasi Bivariat kotak dialog terbuka (Gambar 6.12).



Gambar 6.12 Korelasi Bivariat dengan pilihan koefisien korelasi Spearman

Langkah 3 Pindahkan kedua variabel ke kotak Variabel dan pilih Spearman di bawah Koefisien Korelasi dan Two-tailed di bawah Uji Signifikansi.

Langkah 4 Klik OK untuk menjalankan analisis.

Langkah 5 Output muncul seperti ditunjukkan di bawah ini:

Output

Correlations				
			Grades obtained by students	Family income (Rc)
Spearman's rho	Grades obtained by students	Correlation Coefficient	1.000	-.396
		Sig. (2-tailed)		.257
		N	10	10
	Family income (Rc)	Correlation Coefficient	-.396	1.000
		Sig. (2-tailed)	.257	
		N	10	10

Interpretasi

Dalam output ini, koefisien korelasi Spearman diberikan pada titik perpotongan antara "grade points obtained" (Nilai yang diperoleh) dan "family income" yang $-0,396$, dan Nilai signifikansi ⁸ untuk dua-ekor adalah $0,257$. Karena Nilai signifikansi lebih besar dari $0,05$ ($p > 0,05$), koefisien korelasi tidak signifikan dan disimpulkan bahwa prestasi akademik tidak terkait dengan pendapatan keluarga dalam kumpulan data ini.

KOEFISIEN KORELASI PERINGKAT (RANK) KENDALL

Ini dilambangkan dengan τ (Tau) dan karena itu sering disebut sebagai Koefisien tau Kendall. Uji tau adalah Uji Hipotesis non-parametrik yang menggunakan koefisien untuk menguji ketergantungan statistik. Uji ini dinamai Maurice Kendall, yang mengusulkan dan mengembangkannya. Koefisien korelasi peringkat Kendall mengevaluasi derajat kesamaan antara dua set peringkat yang diberikan ke satu set obyek yang sama. Koefisien ini tergantung pada jumlah inversi pasangan obyek yang diperlukan untuk mengubah satu urutan peringkat ke yang lainnya.

Koefisien Kendall berbeda dari koefisien korelasi Spearman. Secara khusus, hal ini adalah ukuran korelasi peringkat, yaitu, nilai-nilai yang di peringkat oleh masing-masing kuantitas, tetapi konkordansi dan kejanggalan dianggap bukan penyimpangan peringkat. Jika dua variabel acak X dan Y memiliki satu set pengamatan $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, sepasang observasi (x_i, y_i) dan (x_j, y_j) dikatakan sesuai jika peringkat untuk kedua nilai setuju: yaitu, jika keduanya $x_i > x_j$ dan $y_i > y_j$ atau jika keduanya $x_i < x_j$ dan $y_i < y_j$; kalau tidak mereka disebut tidak mirip (discordant).

Rumus berikut ini digunakan untuk menghitung nilai korelasi peringkat Kendall:

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)}$$

dimana,

n_c – jumlah konkordan

n_d – jumlah ketidakmiripan (discordant)

n – jumlah total pasangan pengamatan yang dimungkinkan

KORELASI KENDALL DENGAN SPSS

Contoh 6.4

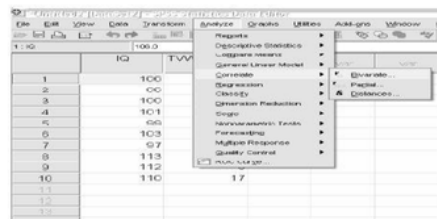
Hitung koefisien korelasi peringkat Kendall untuk data pada IQ 10 orang dan jumlah jam menonton TV

IQ	106	86	100	101	199	103	97	113	112	110
No. of hours of TV watching	7	0	27	50	28	29	20	12	6	17

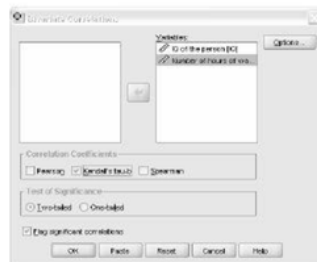
Prosedur untuk korelasi peringkat Kendall akan sama dengan korelasi Pearson di SPSS.

Langkah 1 Beri nama variabel dalam Variabel VIEW dan data di Data Editor.

Langkah 2 Klik Analisis, lalu pilih Korelasi dan kemudian pilih Bivariat, dialog kotak Korelasi Bivariat terbuka (Gambar 6.13).



Gambar 6.13 Memilih Bivariate dari menu utama



Gambar 6.14 Korelasi Bivariat dengan variabel Uji dan Kendall's Tau-b dipilih

Langkah 3 Transfer kedua variabel ke kotak Variabel dan pilih Kendall's Tau-b pada koefisien korelasi dan Two-tailed pada Uji Signifikan.

Langkah 4 Klik OK untuk menjalankan analisis, output muncul seperti yang diberikan di bawah ini

Output

Correlations			
		IQ of the person	Number of hours of watching TV
Kendall's tau_b	IQ of the person	1	-.11
	Correlation Coefficient		.65
	Sig. (2-tailed)		.10
	Number of hours of watching TV	-.11	1
	Correlation Coefficient		.65
	Sig. (2-tailed)		.10

Interpretasi

Pada output ini, Koefisien korelasi Kendall Tau-b diberikan pada titik persilangan antara IQ orang dan jumlah jam menonton TV, yang = -0,11, dan signifikansi untuk dua ekor adalah 0,65. Karena, nilai signifikansi lebih besar dari 0,05 ($p > 0,05$), koefisien korelasi tidak signifikan dan disimpulkan bahwa IQ orang tersebut tidak terkait dengan jumlah jam menonton TV dalam kumpulan data ini.

KORELASI GANDA

Dalam contoh sebelumnya, kita telah membahas hubungan antara dua variabel. Kita sekarang akan memperluas pertimbangan ini untuk hubungan antara tiga atau lebih variabel menggunakan prosedur korelasi ganda. Korelasi ganda (kadang-kadang disebut korelasi regresi berganda atau korelasi linier berganda) merupakan perpanjangan dari korelasi linier. Jika tidak ada variabel yang diasumsikan tergantung secara fungsional satu sama lain, maka kita dapat menerapkan korelasi ganda. Dalam semua kasus, asumsinya adalah bahwa variabel terdistribusi normal. Dalam korelasi sederhana, sampel berasal dari populasi normal bivariat; sedangkan dalam korelasi ganda, sampel (di mana pengamatan atau pengukuran dikumpulkan untuk lebih dari dua variabel) adalah distribusi normal multivariat. Dalam korelasi ganda, semua variabel y , x_1 , x_2 , x_3 , ... dll harus diperlakukan sebagai y .

Keuntungan dari korelasi ganda atas korelasi bivariat sederhana

1. Efek Curvilinear dapat diuji
2. Efek interaksi dapat diuji.

3. Peneliti dapat mempelajari seberapa banyak variasi dalam variabel dependen dijelaskan oleh satu set variabel yang bertentangan dengan yang lain.

4. Kepentingan relatif dari setiap variabel dapat diidentifikasi.

$$R = \frac{\sqrt{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}(r_{yx_2})(r_{x_1x_2})}}{\sqrt{1 - r_{x_1x_2}^2}}$$

Rumus berikut digunakan untuk menghitung korelasi ganda yang melibatkan tiga variabel.

di mana, R = koefisien korelasi ganda, r = koefisien korelasi sederhana dari dua variabel yang diberikan, yaitu antara y dan x_1 , y dan x_2 , dan x_1 dan x_2 dan seterusnya.

Interpretasi

Berbeda dengan koefisien korelasi sederhana r , yang menjelaskan kekuatan dan arah hubungan antar variabel, koefisien korelasi ganda R , hanya menguraikan kekuatan asosiasi. Karena ini merupakan ukuran kekuatan hubungan linear antara variabel y dan himpunan variabel $x_1, x_2, \dots, x_p, \dots$, nilai R tidak pernah negatif. R dapat mengambil nilai ap⁶ pun dari 0 dan +1. Jika nilai R adalah 1, ada hubungan linier sempurna. Jika R sama dengan nol, maka tidak ada hubungan linear antara variabel. Untuk nilai mulai dari 0 hingga 1, signifikansi R diprediksi dengan Uji signifikansi. SPSS memungkinkan peneliti untuk melakukan keduanya secara bersamaan.

KORELASI GANDA DENGAN SPSS

Karena pengerjaan korelasi berganda dan regresi berganda memiliki prosedur yang sama, mengerjakan korelasi berganda sama dengan regresi berganda (Lihat Bab 7).

APAKAH KOEFISIEN KORELASI SEPENUHNYA MENGUNGKAPKAN HUBUNGAN ANTARA VARIABEL?

Tidak ada diskusi tentang korelasi akan lengkap tanpa diskusi sebab-akibat. Dalam analisis korelasi, dimungkinkan untuk menjelaskan hubungan antara dua variabel, tetapi tidak menjelaskan satu variabel sebagai sebab dan yang lain sebagai efek. Misalnya, ada korelasi yang tinggi antara jumlah botol Pepsi yang terjual dan jumlah kematian yang tenggelam di musim panas. Apakah itu berarti bahwa seseorang tidak boleh minum Pepsi sebelum seseorang berenang? Belum tentu. Ini adalah contoh korelasi tanpa sebab-akibat.

Di sisi lain ada korelasi positif yang tinggi antara merokok dan insiden kanker. Perusahaan rokok mengatakan bahwa orang yang merokok lebih gugup dan orang yang gugup lebih rentan terhadap kanker. Sementara dokter mengatakan bahwa merokok memang menyebabkan kanker. Tapi faktanya mereka yang merokok lebih banyak terkena kanker atau merokok menyebabkan kanker. Di sini merokok adalah penyebab dan efek kanker dan karenanya tidak merokok. Ini adalah contoh korelasi sebab-akibat. Dalam contoh jika kita hanya melakukan korelasi, kita hanya dapat menemukan kekuatan hubungan antara variabel dan tidak dapat melanjutkan lebih jauh ke prediksi. Kita bisa pergi untuk prediksi dengan analisis regresi. Bab 7 membahas analisis regresi.

REVIEW LATIHAN

1. Hitung korelasi Karl Pearson antara usia suami dan istri dan gambar diagram pencar.

Age of husband	25	29	30	27	28	40	54	47	35	60
Age of wife	23	27	30	26	26	38	50	44	33	56

2. Gambarkan diagram pencar untuk data berikut pada panjang dan Width dari 10 ikan, hitung koefisien korelasi dan tafsirkan hasil Anda.

Length of fishes (cm)	13	15	15	16	17	19	21	23	16	20
Breadth of fishes (cm)	4	4	4	4	4	6	6	8	4	5

3. Hitung koefisien korelasi Karl Pearson untuk data berikut pada Width dan berat lengan pada bintang laut.

Arm width (in mm)	12	9	8	10	11	13	7
Weight (in g)	14	8	6	9	11	12	3

4. Dua hakim dalam lomba melukis memberi peringkat dua belas mahasiswa sebagai berikut. Hitunglah koefisien korelasi Rank dan nyatakan derajat kesepakatan antara kedua hakim.

Judge 1	5	11	2	3	4	1	6	8	7	10	9	12
Judge 2	4	8	5	2	1	6	7	10	9	11	12	3

5. Hitung koefisien korelasi Kendall untuk data berikut tentang permintaan dan penawaran.

Year	2000	2001	2002	2003	2004	2005	2006	2007	2009	2009
Supply	125	160	164	174	155	170	165	162	172	175
Demand	112	125	192	190	165	174	124	127	152	169

Bab 7

REGRESI

Pada bab sebelumnya, kita membahas tentang korelasi linier sederhana dan korelasi ganda. Kedua prosedur ini memungkinkan peneliti untuk menemukan derajat dan arah atau hanya tingkat hubungan antara dua variabel. Tetapi prosedur ini bertujuan untuk menemukan hubungan fungsional (yaitu, sebab dan akibat hubungan) antara variabel yang diteliti. Hubungan sebab dan akibat antara variabel dipelajari dengan analisis regresi. Istilah "Regresi" pertama kali digunakan oleh Sir Francis Galton, seorang ahli biometrik Inggris. Istilah regresi didefinisikan sebagai ukuran matematika dari hubungan rata-rata antara dua atau lebih variabel dalam kumpulan data.

Dalam sebagian besar kasus, hubungan antara dua variabel dapat merupakan ketergantungan fungsional satu dengan yang lain. Besarnya salah satu variabel (variabel dependen) diasumsikan ditentukan oleh fungsi besarnya variabel kedua (variabel independen) sedangkan sebaliknya tidak mungkin. Misalnya, dalam hubungan antara tekanan darah dan usia pada manusia, tekanan darah dapat dianggap sebagai variabel dependen dan usia sebagai variabel independen. Cukup diasumsikan bahwa besarnya tekanan darah seseorang mungkin merupakan fungsi dari usia. Dalam contoh ini, usia bukan satu-satunya penentu biologis tekanan darah, tetapi dianggap sebagai salah satu faktor penentu. Untuk alasan ini, variabel independen disebut prediktor, atau variabel regresi dan variabel dependen disebut sebagai respons, atau variabel kriteria.

REGRESI LINEAR SEDERHANA

Istilah regresi linier sederhana mengacu pada fakta bahwa hanya dua variabel yang dipertimbangkan. Data analisis regresi sederhana terdiri dari variabel dependen dan variabel independen. Hubungan ini dipelajari secara matematis dan grafis.

PERSAMAAN REGRESI LINEAR SEDERHANA

Hubungan fungsional sederhana dari satu variabel ke variabel lainnya dalam suatu populasi adalah regresi linier sederhana. Hubungan semacam ini diungkapkan dalam bentuk persamaan:

$$Y_i = a + b X_i$$

dimana,

Y_i - item dari variabel Y (tergantung)

X_i - ith item variabel X (independen),

a - konstan dan disebut Y cegatan (intercept), yaitu, Value Y pada garis regresi ketika $X = 0$

b -parameter disebut sebagai koefisien regresi atau kemiringan (perubahan dalam y dengan perubahan satu unit dalam X).

Value b secara teoritis dapat berkisar dari $-\infty$ to $+\infty$, termasuk nol.

Dari hubungan fungsional persamaan $Y_i = a + b X_i$, kita dapat memprediksi Value Y yang paling mungkin untuk Value X yang diberikan. Untuk melakukan ini kita perlu mencari konstanta a dan parameter b dalam persamaan. Konstanta a dan b dihitung dari rumus berikut.

$$a = \bar{X} - b\bar{Y}$$
$$b = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}}$$

dimana,

\bar{X} - rata-rata variable X,

\bar{Y} - rata-rata variable Y dan

Σ - penjumlahan

Ketika kita memformulasikan persamaan regresi akan dapat memprediksi Value Y yang paling mungkin untuk Value X yang diberikan.

DIAGRAM PENCAR DAN GARIS PALING PAS (BEST FIT)

Seperti pada bab sebelumnya, kita dapat memplot variabel Y dan X pada lembar grafik menggunakan ordinat (sumbu-Y) untuk variabel dependen (Y) dan absis (sumbu-X) untuk variabel independen (X). Sepasang data X dan Y dapat dilambangkan sebagai (X_1, Y_1) , yang lain sebagai (X_2, Y_2) , yang lain sebagai (X_3, Y_3) , dll. Dengan demikian data muncul sebagai titik pencar, setiap titik mewakili sepasang X dan nilai Y. Representasi grafik yang dihasilkan disebut diagram pencar. Ini memberikan penyebaran variabel yang menunjukkan sifat hubungan. Garis yang ditarik melalui titik-titik ini sedemikian rupa sehingga penyimpangan titik pencar dari garis adalah yang terkecil. Garis seperti itu disebut garis paling pas.

Perhatikan data dalam Contoh 1 di Bab 6 di mana ada ⁶ dua variabel yaitu usia, variabel independen dan tekanan darah, variabel dependen. Dari sebaran data ini, tampak bahwa pengukuran tekanan darah berhubungan linier dengan usia. Ketika kita menggambar kurva "eye fit" akan ada variabilitas data yang cukup besar di sekitar garis lurus yang mungkin kita tarik. Oleh karena itu, kita berusaha untuk menggambarkan apa yang biasa disebut garis "paling pas" melalui data. Garis yang paling pas tersebut disebut "garis regresi". Kriteria untuk "kesesuaian" yang umumnya digunakan menggunakan konsep kuadrat terkecil. Setiap nilai X akan memiliki nilai Y yang sesuai pada garis yang mungkin kita tarik melalui titik-titik data yang tersebar. Nilai Y ini diwakili untuk membedakannya dari nilai Y yang benar-benar diamati dalam sampel.

Kriteria kuadrat terkecil mempertimbangkan deviasi vertikal dari setiap titik dari garis (yaitu, deviasi yang digambarkan tersebut), dan mendefinisikan garis paling pas yang menghasilkan Nilai terkecil untuk jumlah kuadrat dari penyimpangan ini untuk semua Nilai Y_i dan \bar{Y} . Artinya, harus minimum, di mana n adalah jumlah titik data yang terdiri dari sampel. Jumlah kuadrat dari penyimpangan ini disebut jumlah sisa kuadrat (atau, kadang-kadang, jumlah kesalahan kuadrat). Dengan demikian ada hubungan fungsional antara garis regresi dan persamaan regresi. Yang terakhir adalah ekspresi matematis dari yang pertama.

REGRESI DAN ONE-WAY ANOVA

Persamaan regresi linier sederhana didasarkan pada a dan b yang perhitungannya melibatkan konsep kuadrat terkecil. Untuk alasan ini salah satu cara analisis varians dilakukan bersamaan dengan analisis regresi. Prosedur ini memungkinkan peneliti untuk mengetahui seberapa jauh varians dalam variabel independen berkontribusi terhadap perubahan dalam variabel dependen.

REGRESI SEDERHANA DENGAN SPSS

Seperti halnya analisis korelasi, SPSS memungkinkan kita untuk melakukan analisis dengan cepat. Mari kita hitung konstanta a dan parameter b untuk pasangan variabel, usia (variabel independen) dan tekanan sistolik (variabel dependen)

Contoh 7.1

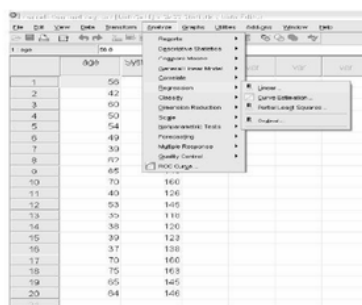
Rumuskan persamaan regresi Y (tekanan sistolik mm Hg) pada X (usia) pada manusia. Prediksi nilai yang paling mungkin dari tekanan sistolik untuk usia 51 dan 68 tahun.

Age	56	42	60	50	54	49	39	62	65	70	40	53	35	38	39	37	70	75	65	64
BP	160	130	125	135	145	115	140	120	140	160	126	145	118	120	123	138	160	163	145	146

Langkah 1 Buka Data Editor dan klik Variabel view. Masukkan nama variabel dan rincian variabel.

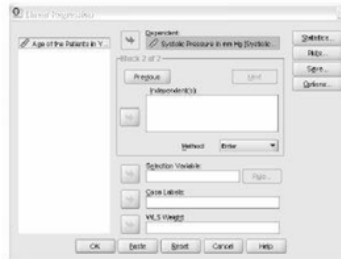
Langkah 2 Klik Data view dan masukkan data di bawah variabel yang sesuai.

Langkah 3 Pilih Analisis dari menu utama dan pilih Regresi dan Linear dari menu seperti pada Gambar 7.1.



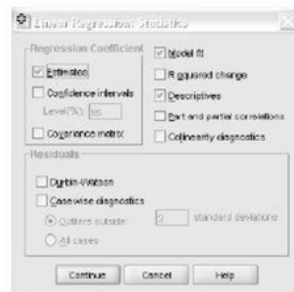
Gambar 7.1 Memilih Regresi Linear dari menu utama

Langkah 4 Dialog Linear Regression terbuka. Pindahkan Tekanan sistolik ke kotak Dependen dan Age ke kotak Independen (Gambar 7.2).



Gambar 7.2 Linear Regression kotak dialog untuk mentrans variabel

Langkah 5 Klik Statistik untuk membuka Linear Regression: Statistik kotak dialog, pilih **Estimates**, **Descriptives** dan **Modul Fit**, lalu klik Lanjutkan (Gambar 7.3).



Gambar 7.3 Linear Regression: dengan opsi **Estimates**, **Descriptives** dan **Modul Fit**

Langkah 6 Klik Plot untuk membuka **Linear Regression: Plot** dan pindahkan ***ZRESID** (dependen atau kriteria) ke kotak Y: dan ***ZPRED** (independen atau prediktor) ke kotak X: (Gambar 7.4) lalu klik Lanjutkan.

Langkah 7 Klik OK untuk menjalankan analisis.



Gambar 7.4 Plot Linear Regression untuk mentransfer variabel Dependen dan Independen

Langkah 8 Output muncul di bawah judul yang berbeda sebagaimana diberikan dalam Output 1, 2, 3, dan 4.

Output 1

Descriptive Statistics			
	Mean	Std. Deviation	N
Systolic Pressure in mm Hg	137.70	15.359	20
Age of the Patients in Years	53.15	12.820	20

Output 2

Coefficients ^a					
Model		Unstandardized Coefficients	Standardized Coefficients	t	Sig.
1	(Constant)	95.625		8.26	.000
	Age of the Patients in Years	.792	0.66	3.73	.002

a. Dependent Variable: Systolic Pressure in mm Hg

Output 3

ANOVA ^b						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1956.827	1	1956.827	13.948	.002 ^a
	Residual	2525.373	18	140.299		
	Total	4482.200	19			

a. Predictors: (Constant), Age of the Patients in Years

b. Dependent Variable: Systolic Pressure in mm Hg

Output 4

Residuals Statistics ^a					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	123.33	155.00	137.70	10.15	20
Residual	-24.71	20.04	1.48E-14	11.53	20
Std. Predicted Value	-1.42	1.7	-3.83E-16	1	20
Std. Residual	-2.09	1.69	1.26E-15	0.97	20

a. Dependent Variable: Systolic Pressure in mm Hg

Interpretasi

Tergantung pada kebutuhan peneliti, nilai-nilai dalam output dapat digunakan untuk interpretasi.

1. Rata-rata dan standar deviasi dari variabel diberikan di bawah Statistik Deskriptif (Output 1). Usia rata-rata orang dalam sampel adalah 53,15 dengan standar deviasi dari 12,820 (dalam makalah penelitian ditulis sebagai $53,15 \pm 12,820$ tahun) dan tekanan sistolik rata-rata adalah 137,70 dengan standar deviasi 15,359 ($137,70 \pm 15,359$ mmHg).

2. Konstanta dan parameter yang diperlukan untuk perumusan persamaan regresi $Y_i = a + bX_i$ diberikan dalam output 2 pada judul koefisien. Nilai yang diberikan pada kolom B terhadap Konstanta adalah Nilai-a (95,625) dan terhadap Usia pasien dalam Tahun adalah Nilai b (0,792).

Oleh karena itu, persamaan regresi $Y_i = a + bX_i$ dirumuskan sebagai: $Y_i = 95,625 + 0,792X_i$

Untuk memprediksi tekanan sistolik yang paling mungkin untuk usia 51 dan 68 tahun:

i. Tekanan sistolik untuk seseorang usia 51 tahun Disini persamaannya diformulasikan sebagai $Y_i = 95,625 + 0,792 \times 51$

$$Y_i = 95,625 + 40,392$$

$$Y_i = 136,017 \text{ mm Hg.}$$

ii. Tekanan sistolik untuk seseorang yang berusia 68 tahun Disini persamaannya diformulasikan sebagai:

$$Y_i = 95,625 + 0,792 \times 68$$

$$Y_i = 95,625 + 53,856$$

$$Y_i = 149.481 \text{ mmHg.}$$

3. Output 3 memberikan hasil pada analisis varians. F-ratio yang diberikan di bawah kolom F adalah 13,948, dan p-value, 0,002 diberikan di bawah kolom Sig. Karena p-value kurang dari 0,01, itu berarti bahwa koefisien regresi dihitung signifikan dan varians dalam variabel independen berkontribusi terhadap perubahan dalam variabel dependen. Oleh karena itu, disimpulkan



Gambar 7.7 Create Scatterplot dengan opsi Fit terpilih

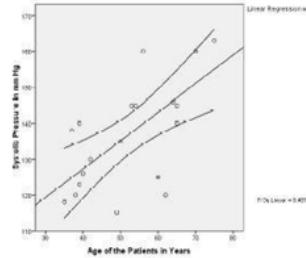
Langkah 4 Ketika Anda memilih Regresi, semua kotak dalam kotak dialog itu akan disorot, sekarang pilih **Constant in Equation** dan Rata-rata hitung di bawah garis Prediksi dan klik OK (Gambar 7.8).



Gambar 7.8 Create Scatterplot dengan opsi yang dipilih **Method, Prediction line and Fit lines for**

Langkah 5 Hasilnya muncul dalam output seperti pada Output 1. Garis tengah adalah garis regresi, yaitu, garis "fit terbaik" dan dua garis lainnya, satu di bawah dan yang lain di atas garis regresi memberikan batas kepercayaan (95%). Titik-titik pencar adalah plot setiap pasangan variabel untuk satu individu.

Output 1



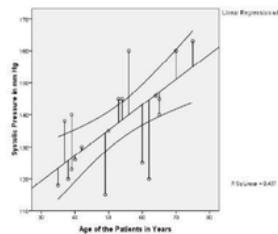
Langkah 6 Untuk menunjukkan deviasi berbagai titik dari garis regresi atau garis klik **Spike** pada **Create Scatterplot** dan pilih **Fit Line** pada **Spike** (Gambar 8.9).



Gambar 7.9 **Create Scatterplot** dengan "Fit Line" dipilih pada **Spikes**

Langkah 7 Klik **OK** untuk mendapatkan output seperti Output 2.

Output 2



Pada Output 2, garis vertikal dari garis yang paling pas dengan titik pencar menunjukkan deviasi.

ANALISIS REGRESI GANDA

10

Tujuan umum dari regresi berganda adalah untuk mempelajari lebih lanjut tentang hubungan antara beberapa variabel independen atau

variabel prediktor dan variabel dependen atau kriteria. Dalam penelitian ilmu sosial dan biologi, prosedur regresi berganda sangat banyak digunakan. Misalnya, peneliti pendidikan mungkin ingin mempelajari prediktor terbaik kesuksesan di perguruan tinggi. Ahli biologi mungkin ingin menentukan prediktor terbaik untuk kelangsungan hidup ikan dalam air yang tercemar. Psikolog mungkin ingin menentukan variabel kepribadian mana yang paling tepat memprediksi penyesuaian sosial. Sosiolog mungkin ingin mengetahui mana dari beberapa indikator sosial yang paling tepat memprediksi apakah kelompok imigran baru akan beradaptasi dengan situasi baru.

Regresi linear sederhana untuk variabel berpasangan dalam suatu populasi adalah

$$Y_i = a + bX_i$$

Dalam hubungan ini, Y dan X masing-masing merupakan variabel dependen dan independen, b adalah koefisien regresi dalam populasi, sebuah (Y intersep) adalah Nilai Y ketika X = nol.

Namun dalam banyak situasi, Y dapat dianggap tergantung pada lebih dari satu variabel. Kemudian,

$$Y_j = a + b_1 X_{1j} + b_2 X_{2j}$$

Di sini, satu variabel (Y) secara linear bergantung pada variabel kedua (X1) dan Y juga secara linear bergantung pada variabel ketiga (X2). Dalam modul regresi berganda ini kita memiliki satu variabel dependen dan dua variabel independen. Dua parameter populasi b1 dan b2, disebut sebagai koefisien regresi parsial; b1 menyatakan berapa banyak Y akan berubah untuk perubahan satuan di X1, jika X2 dipertahankan konstan. Kadang-kadang dikatakan bahwa b1 adalah ukuran hubungan Y ke X1 setelah menghapus efek X2. Demikian pula, b2 menggambarkan tingkat perubahan Y sebagai perubahan X2, dengan X1 konstan. b1 dan b2 disebut koefisien regresi parsial, karena masing-masing hanya mengungkapkan sebagian dari hubungan ketergantungan. Intersep Y, a (kadang-kadang ditetapkan sebagai b0) adalah nilai Y ketika keduanya X1 dan X2 adalah nol.

Jika kita sampel populasi yang mengandung tiga variabel (Y, X1 dan X2) persamaan regresi berganda dapat dinyatakan sebagai

$$Y_j = a + b_1 X_{1j} + b_2 X_{2j}$$

di mana Y_j = variabel dependen, a adalah konstanta yang disebut intersep Y di mana nilai variabel independen X_1 dan X_2 adalah 0, b_1 slope (kemiringan) X_1 dengan X_2 konstan, b_2 = slope X_2 dengan X_1 konstan.

Catatan: Deskripsi dan rumus di atas dijelaskan untuk membuat peneliti memiliki pemahaman yang lebih baik tentang prinsip-prinsip melakukan regresi berganda dan memungkinkan mereka untuk menginterpretasikan hasil.

REGRESI GANDA DENGAN SPSS

Contoh 7.3

Catatan rumah sakit memberikan data volume udara expiratory paksa (liter), kapasitas vital (liter) dan kapasitas paru total (liter). Volume udara ekspiratori paru paksa tergantung pada kapasitas vital dan kapasitas paru total. Hitung koefisien regresi berganda dan rumuskan persamaan regresi berganda.

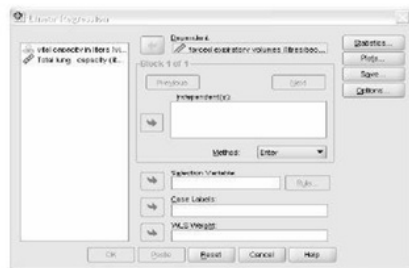
Dalam contoh ini hanya tiga variabel yang diberikan. Dari ketiganya, volume udara ekspiratori paksa adalah variabel dependen dan kapasitas vital dan kapasitas paru total adalah variabel independen.

Vital capacity (litre)	2.2	1.5	1.6	3.4	2.0	1.9	2.2	3.3	2.4	0.9	1.0	2.1	2.8	1.5	1.6
Total lung capacity (litre)	2.5	3.2	4.0	4.4	4.4	3.3	3.2	3.3	3.7	2.8	3.6	3.6	4.3	3.2	5.0
Expiratory air volume (litre)	1.0	1.0	1.4	2.6	1.2	1.5	1.6	1.6	2.1	0.7	0.7	1.1	2.5	1.0	1.4

Langkah 1 ¹³ Masukkan nama variabel dalam Variabel view dan data dalam Data editor.

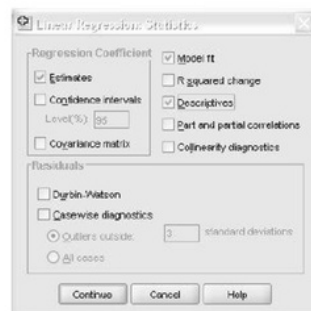
Langkah 2 Pilih Analisis dari menu utama, kemudian Regresi dan kemudian pilih Linear untuk membuka Kotak dialog regresi linier.

Langkah 3 Pindahkan volume udara ekspiratori paksa ke kotak Dependent (Gambar 7.10) dan kapasitas Vital dan kapasitas paru total ke kotak Independen.



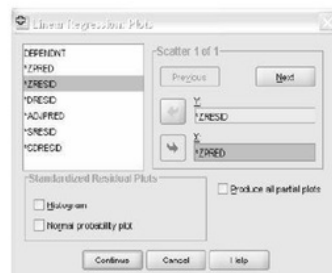
Gambar 7.10 Linear Regression untuk memindahkan variabel Dependen dan Independen

Langkah 4 Klik Statistik untuk membuka Linear Regression: Pilih Estimates, Model Fit, dan Deskriptif lalu klik Lanjutkan (Gambar 7.11).



Gambar 7.11 Linear Regression memilih Estimates, Model Fit dan Descriptives

Langkah 5 Klik Plot untuk membuka Plot Linear Regression dan pindahkan *ZRESID (variabel dependen atau kriteria) ke kotak Y dan *ZPRED (variabel independen atau prediktor) ke kotak X (Gambar 7.12) lalu klik Lanjutkan.



Gambar 7.12 Plot Linear Regression untuk mengatur transfer variabel Y dan X

Langkah 6 Klik OK untuk menjalankan analisis, output (1, 2, 3, 4 dan 5) muncul.

Output 1

Descriptive Statistics			
	Mean	Std. Deviation	N
forced expiratory volumes (litres/second)	1.428	.5844	15
vital capacity in liters	2.027	.7353	15
Total lung capacity (liters)	3.633	.6758	15

Output 2

Model Summary ^a				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.870 ^a	.758	.717	.3108

a. Predictors: (Constant), Total lung capacity (liters), vital capacity in liters

b. Dependent Variable: forced expiratory volumes (litres/second)

Output 3

ANOVA ^b						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.622	2	1.811	18.753	.000 ^a
	Residual	1.159	12	.097		
	Total	4.781	14			

a. Predictors: (Constant), Total lung capacity (liters), vital capacity in liters

b. Dependent Variable: forced expiratory volumes (litres/second)

Output 4

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.622	2	1.811	18.753	.000 ^a
	Residual	1.159	12	.097		
	Total	4.781	14			

a. Predictors: (Constant), Total lung capacity (liters), vital capacity in liters

b. Dependent Variable: forced expiratory volumes (litres/second)

Output 5

Residuals Statistics ^a					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	0.53	2.45	1.43	.5086	15
Residual	-.4714	.4360	.0000	.2877	15
Std. Predicted Value	-1.76	2	1.85E-16	1	15
Std. Residual	-1.52	1.4	-.74E-17	0.93	15

a. Dependent Variable: forced expiratory volumes (litres/second)

Interpretasi

1. Output 1 memberikan statistik deskriptif yaitu rata-rata dan standar deviasi untuk tiga variabel dan jumlah variasi. Volume ekspiratori rata-rata adalah 1,428 0,584 (liter/detik), kapasitas vital adalah 2,027 0,735 (liter) dan kapasitas paru-paru adalah 3,633 0,676 (liter).

2. Pada Output 2, Model Ringkasan, nilai 0,870 yang diberikan di bawah kolom R adalah koefisien korelasi berganda. Ketiga variabel ini berkorelasi secara signifikan.

3. Output 3 memberikan hasil ANOVA, karena p-value yang diberikan di bawah kolom signifikansi adalah <0,01. Ini berarti bahwa varians dalam kapasitas vital dan kapasitas paru total (dua variabel independen atau variabel prediktor) berkontribusi secara signifikan terhadap perubahan volume udara ekspiratori paksa (variabel dependen).

4. Output 4 memberikan koefisien; Nilai-nilai ini diperlukan untuk memformulasikan persamaan regresi. Nilai di bawah kolom B terhadap konstanta adalah nilai "a" (Y-intercept) dalam persamaan regresi dan nilai-nilai terhadap kapasitas paru total (0,288) dan kapasitas vital (0,581) menentukan slope garis regresi dan merupakan nilai dari b1 dan b2 dalam persamaan regresi berganda:

$$Y_j = a + b_1 X_{1j} + b_2 X_{2j}$$

Oleh karena itu, persamaan regresi berganda diformulasikan sebagai

$$Y_j = -0,796 + 0,288 X_{1j} + 0,581 X_{2j}$$

5. Output 5 memberikan statistik residual di mana jumlah kasus, rata-rata dan standar deviasi untuk nilai prediksi variabel dependen diberikan.

ANALISIS KORELASI VS ANALISIS REGRESI

1. Dalam analisis korelasi, kita mengidentifikasi kekuatan dan arah dari hubungan linear antara dua variabel acak. Ini hanya menjelaskan seberapa besar ρ satu variabel cenderung berubah ketika yang lain berubah. Analisis regresi bertujuan untuk membangun hubungan fungsional antara dua variabel yang diteliti dan dengan menggunakan hubungan ini nilai satu variabel diprediksi atau diperkirakan untuk setiap nilai yang diberikan dari variabel lain.

2. Dalam korelasi, tidak masalah mana dari dua variabel disebut "X" atau "Y". Kita akan mendapatkan koefisien korelasi yang sama bahkan jika kita mengubah dua variabel yaitu, mempertimbangkan X sebagai Y.

3. Dalam regresi, dua variabel 'X' dan 'Y' dianggap masing-masing sebagai variabel 'independen atau prediktor' dan 'dependen atau kriteria'. Kita tidak dapat mengubah variabel, jika kita mengubah dua variabel, maka akan mendapatkan hasil yang berbeda.

4. Koefisien korelasi r adalah ukuran relatif dari hubungan linear antara variabel yang diteliti dan merupakan bilangan murni yang terletak antara ± 1 . Di sisi lain, koefisien regresi b_{yx} adalah ukuran absolut yang mewakili perubahan nilai variabel Y untuk perubahan unit dalam nilai variabel X.

5. Korelasi tidak mempelajari hubungan 'sebab' dan 'efek'. Hanya mengukur seberapa baik dua variabel berhubungan satu sama lain. Dalam regresi kita mempelajari hubungan sebab dan akibat dan kita dapat memprediksi nilai yang paling mungkin dari variabel dependen untuk variabel independen yang diberikan.

6. Dalam korelasi kita tidak dapat memanipulasi variabel secara eksperimental. Dengan regresi linier, variabel X adalah sesuatu yang sering kita manipulasi secara eksperimental (waktu, konsentrasi ...) dan efek pada variabel Y diukur.

REVIEW LATIHAN

1. Rumuskan persamaan regresi untuk tinggi putra dan ayah untuk data berikut:

Hitung tinggi badan anak yang paling mungkin ketika tinggi ayah adalah 166 dan 172. Gambarlah diagram pencar dan garis yang paling pas.

Height of father (cm)	165	160	157	158	168	170	171	169	165	163
Height of son (cm)	160	162	155	156	165	168	170	159	161	164

2. Gambar diagram pencar. Rumuskan persamaan regresi berat pada tinggi untuk data tekanan oksigen parsial (mmHg) dalam air dan oksigen yang dikonsumsi (mg/jam) oleh ikan.

Partial pressure of oxygen (mmHg)	117	115	109	105	100	95	93	89	85	75
Oxygen consumed (mg)	6.9	6.9	6.7	6.5	6.4	5.0	5.5	3.3	2.3	1.9

3. Berikut adalah data menunjukkan konsentrasi aldrin pestisida (mg/l) dan kematian ikan dalam %. Gambarlah diagram pencar dan garis paling pas dengan spikes. Rumuskan persamaan regresi kematian pada konsentrasi pestisida.

Concentration of pesticide aldrin (mg/l)	0.01	0.65	0.30	0.50	0.70	0.40	0.80	1.0	0.9
Mortality (%)	6	72	16	48	65	36	78	98	91

4. Rumuskan persamaan regresi untuk data berikut tentang penawaran dan harga.

Supply	80	82	91	83	85	89	96	93	90	92
Price	146	140	130	117	133	127	115	95	100	97

5. Rumuskan persamaan regresi berganda untuk data berikut pada variabel X, Y dan Z.

X	112	125	118	118	121	125	131	135	140	145
Y	106	102	102	104	98	96	90	85	86	82
Z	115	125	192	193	168	145	150	155	156	157

Bab 8

UJI CHI-SQUARE

Karl Pearson pada tahun 1900 mengembangkan prosedur statistik untuk menguji signifikansi perbedaan antara nilai-nilai eksperimental dan nilai-nilai teoritis yang diperoleh di beberapa teori atau Hipotesis. Uji ini dikenal sebagai χ^2 Uji Chi-square (Kai kuadrat). Ini adalah Uji "kesesuaian paling pas" (goodness of fit) dan digunakan untuk mengetahui apakah penyimpangan antara observasi (eksperimen) dan teori mungkin disebabkan secara kebetulan (fluktuasi dalam sampling) atau faktor lainnya. Uji chi-square berlaku untuk data yang diperoleh dengan cara pencacahan (enumerasi) dan bukan untuk data yang dikumpulkan dengan pengukuran pada skala kontinyu (variabel kontinyu kuantitatif).

Rumus untuk menghitung Nilai chi-square adalah:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

dimana,

Σ -penjumlahan

O_i -frekuensi observasi

E_i -frekuensi diharapkan.

PERSYARATAN DATA UNTUK UJI CHI-SQUARE

i. **Variabel kualitatif atau kategoris.** Ada sejumlah variabel kualitatif dalam populasi manusia seperti warna kulit, mata, rambut, dll. Individu yang memiliki karakter ini dihitung dan jumlah kejadian dalam sampel diberikan dalam bentuk frekuensi (disebut frekuensi yang diamati). Karakteristik ini diharapkan terjadi pada frekuensi tertentu, yang disebut sebagai frekuensi

yang diharapkan. Demikian pula, ada sejumlah variabel kategori seperti golongan darah, tingkat pendidikan, tingkat pekerjaan, dll. Variabel-variabel ini dikategorikan dan jumlah individu yang memiliki karakter ini dihitung dan jumlah kejadian dalam sampel diberikan dalam bentuk frekuensi terhadap masing-masing kategori. Misalnya, jumlah orang dengan golongan darah yang sama dihitung dan diberikan terhadap setiap kategori seperti A, B, AB dan

ii. **Variabel kualitatif atau kategori dalam tabel kontingensi.** Kadang-kadang peneliti mungkin tertarik untuk mengetahui apakah dua karakter kualitatif cenderung terjadi bersama-sama dalam suatu populasi tertentu. Dalam kasus ini, individu dihitung untuk ada atau tidaknya dua karakter spesifik dan diberikan dalam bentuk tabel kontingensi.

Hair colour	Eye colour	
	Black	Brown
Black	50	10
Brown	9	51

Seorang pencinta lingkungan mungkin tertarik untuk mengetahui apakah dua spesies cenderung terjadi bersama dalam suatu ekosistem tertentu. Seorang farmakolog mungkin tertarik untuk mengetahui apakah obat baru efektif dalam menyembuhkan penyakit, seorang dokter mungkin tertarik untuk menemukan hubungan antara dua kondisi yang sakit pada pasien. Dalam semua kasus ini, data disajikan dalam bentuk tabel dengan variabel, masing-masing dalam baris dan kolom dan di bawah setiap kategori pengamatan ditempatkan dalam dua kategori. Tabel ini adalah tabel kontingensi 2×2 , karena data disajikan dalam dua baris dan dua kolom. Bergantung pada jumlah kategori untuk setiap variabel, jumlah baris dan kolom bervariasi.

iii. **Variabel kualitatif atau kategori dalam bentuk proporsi.** Seorang demografer mungkin tertarik untuk mengetahui apakah proporsi status perkawinan laki-laki di empat kota berbeda adalah sama. Ia menghitung jumlah orang – menikah dan lajang dan menyajikannya dalam bentuk proporsi seperti menikah/Total dan lajang/Total di setiap kota.

Kondisi untuk Uji Validitas Chi-square

Uji chi-square hanya dapat digunakan jika kondisi berikut dipenuhi:

- i. N , frekuensi total harus cukup besar, lebih dari 50.
- ii. Pengamatan dalam sampel harus independen. Ini berarti tidak ada item yang harus dimasukkan dua kali atau lebih dalam sampel.
- iii. Lebih disukai, setiap frekuensi teoritis harus lebih besar dari 10 tetapi tidak boleh kurang dari 5.
- iv. Data harus diberikan dalam unit asli.

APLIKASI UJI CHI-SQUARE

i. Uji Chi-square "goodness of fit" atau dengan hipotesis apriori. Kesesuaian paling pas adalah istilah χ^2 yang digunakan untuk menunjukkan seberapa jauh distribusi frekuensi yang diamati sesuai baik dengan distribusi frekuensi yang diharapkan berdasar pada beberapa teori atau harapan. Karena frekuensi yang diharapkan dihitung berdasarkan beberapa teori atau hipotesis, juga disebut Uji Chi-square dengan Hipotesis apriori.

7
Dalam hal ini Hipotesis nol (H_0), yaitu: "Tidak ada perbedaan antara frekuensi yang diamati dan frekuensi yang diharapkan". Hipotesis alternatif (H_A) : "Ada perbedaan antara frekuensi yang diamati dan frekuensi yang diharapkan".

ii. Uji chi-square untuk hubungan antara atribut atau independensi atribut atau Uji chi-square tanpa Hipotesis apriori. Ini digunakan untuk menemukan apakah dua karakter memiliki kecenderungan untuk tetap bersama atau tetap independen. Misalnya, apakah seorang pasien yang menderita penyakit disembuhkan karena obat atau pemulihan dari penyakit tidak tergantung pada obat. Dalam hal ini, frekuensi diberikan dalam bentuk tabel kontingensi.

Dengan demikian, Hipotesis nol (H_0) adalah "Tidak ada hubungan antara dua karakter", yaitu obat tidak efektif dalam menyembuhkan penyakit. Hipotesis alternatif adalah "Ada hubungan antara dua karakter". yaitu, obat efektif dalam menyembuhkan penyakit.

iii. *Uji chi-square untuk menguji kesetaraan proporsi.* Selain Uji chi-square di atas memungkinkan seseorang untuk mengetahui signifikansi perbedaan antara dua atau lebih proporsi populasi. Sebagaimana dinyatakan di atas, jika peneliti tertarik untuk mengetahui apakah proporsi pria yang sudah menikah sama di empat kota besar. Hipotesis nol (H_0): "Proporsi pria yang sudah menikah adalah sama di empat kota besar. Hipotesis alternatif (H_A) adalah "Proporsi pria yang sudah menikah tidak sama di empat kota besar".

iv. *Uji chi-square homogenitas.* Kadang-kadang kita perlu membandingkan sampel yang diperoleh dari dua populasi yang berbeda dalam beberapa karakter. Sebagai contoh, sampel diperoleh dari dua populasi yaitu "normal" dan "pasien penyakit hepatitis"; kedua sampel ini kemudian diklasifikasikan menjadi dua:

1. orang dengan hepatitis dan
2. orang tanpa hepatitis

Di sini Chi-square digunakan untuk mengetahui apakah dua populasi dari sampel diperoleh adalah homogen atau tidak. Dalam hal ini Hipotesis nol (H_0) adalah "Dua populasi sampel yang diperoleh adalah homogen". Hipotesis alternatif adalah "Dua populasi sampel tersebut tidak homogen".

Meskipun, empat aplikasi yang berbeda diberikan untuk pengujian, semua uji ini pada dasarnya sama dengan uji untuk menemukan apakah frekuensi yang diamati sesuai dengan frekuensi yang diharapkan, yaitu, Hipotesis nol (H_0) = $O_i - E_i = 0$ hanya pernyataan Hipotesis nol berbeda .

Prosedur Untuk Melakukan Uji Chi-Square

- i. Buat Hipotesis Nol dan Alternatif
- ii. Hitung frekuensi yang diharapkan
- iii. Hitung deviasi antara frekuensi yang diharapkan dan yang diamati
- iv. Kuadratkan deviasi dan bagi dengan masing-masing frekuensi yang diharapkan
- v. Substitusikan ke dalam rumus

vi. Sebutkan derajat kebebasan χ^2

vii. Merujuk pada nilai uji yang ditabulasikan untuk tingkat kebebasan pada $p=0,05$ (atau $0,01$)

viii. Bandingkan Nilai chi-kuadrat yang dihitung dengan Nilai chi-square pada tabel

ix. Tafsirkan hasilnya.

Prosedur di atas dilakukan secara rutin untuk perhitungan secara manual atau dengan kalkulator. Tetapi ketika kita menggunakan SPSS, tidak perlu melakukan salah satu langkah dari langkah ii hingga viii, kita hanya perlu membuat Hipotesis, memasukkan data dengan benar, menjalankan analisis dan menginterpretasikan hasilnya.

50

UJI CHI-SQUARE DENGAN SPSS

Uji Chi-square untuk "Goodness of Fit" dengan SPSS

Melalui uji ini memungkinkan kita untuk menguji:

i. apakah frekuensi yang diamati untuk variabel kategori berbeda dengan frekuensi yang diharapkan berdasarkan beberapa Hipotesis atau teori, dan

ii. apakah frekuensi yang diamati dari suatu variabel tertentu mengikuti distribusi teoritis seperti distribusi Binomial atau Poisson.

"Goodness of Fit" untuk Variabel Kategorikal

Contoh 8.1

Dipercaya bahwa populasi manusia secara umum di suatu wilayah terdiri dari 10% Hispanik, 10% Asia, 10% Afrika-Amerika, dan 70% rakyat kulit putih. Dalam survei sampel, tercatat ada 24 orang Hispanik, 11 orang Asia, 20 orang Afrika-Amerika, dan 145 orang kulit putih. Ujilah apakah frekuensi yang diamati berbeda secara signifikan dari nilai yang dihipotesiskan.

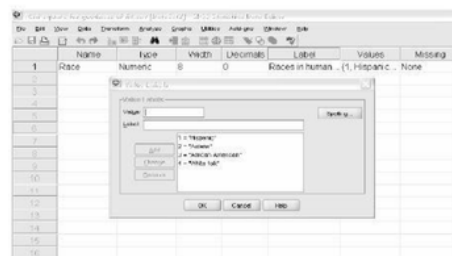
(Rasio yang diharapkan untuk kategori ras manusia adalah = 10: 10: 10: 70) Hipotesis:

H0: Tidak ada perbedaan antara frekuensi yang diamati dan yang diharapkan.

HA: Ada perbedaan antara frekuensi yang diamati dan yang diharapkan.

Langkah 1 Buka file data baru di SPSS.

Langkah 2 Beri nama variabel dalam Variabel view. Klik Variabel View dan beri nama variabel di baris pertama sebagai "Ras", pilih numerik pada Type, pilih 0 di Desimal, label variabel di kolom Label: "Ras populasi manusia" (Ini akan muncul di output). Di kolom Nilai, klik area abu-abu untuk Nilai Label dan tetapkan nomor Nilai dan label sebagai 1 untuk Hispanik, 2 untuk Asia, 3 untuk Afrika-Amerika, dan 4 untuk Putih dan klik OK (Gambar 8.1).



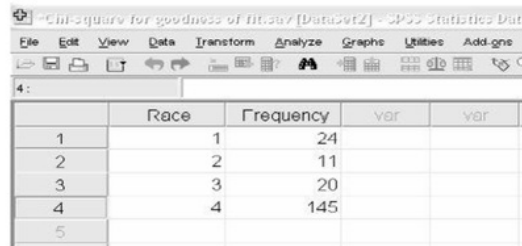
Gambar 8.1 Tampilan Variabel dengan kotak dialog Label Value

Ketik Frekuensi di baris kedua Namr, pilih Numerik pada Type, pilih 0 di Desimal, label variabel di kolom Label: "Frekuensi berbagai ras berbeda" (Ini akan muncul di output). Tidak perlu melabeli kolom Value (Gambar 8.2).

	Name	Type	Width	Decimals	Label	Values	Missing
1	Race	Numeric	8	0	Races in human... (1, Hispanic... None		
2	Frequency	Numeric	8	2	Frequency	None	None
3							
4							
5							

Gambar 8.2 Tampilan Variabel dengan nama dan frekuensi variabel

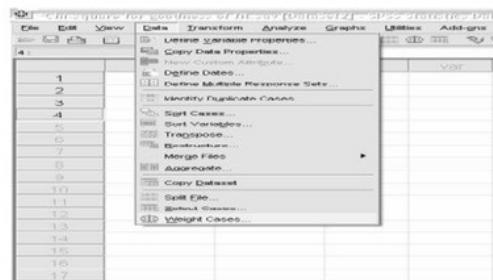
Langkah 3 Klik **16**ta view untuk beralih ke tampilan data. File tersebut muncul seperti pada Gambar 8.3 dan masukkan data.



	Race	Frequency	var	var
1	1	24		
2	2	11		
3	3	20		
4	4	145		
5				

Gambar 8.3 Tampilan Data dengan data yang dimasukkan

Langkah 4 Klik Data di menu utama, lalu klik **Weight Cases** (Gambar 8.4).



Gambar 8.4 Memilih Weight Cases dari menu utama

Ini membuka kotak dialog Weight Cases (Gambar 8.5).



Gambar 8.5 Kotak dialog Weight Cases

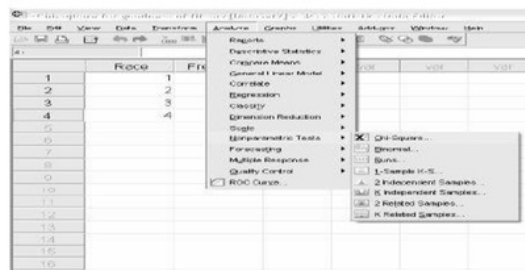
Langkah 5 Pilih **Weight Cases by**, klik pada nama "Frekuensi berbagai ras berbeda" dan pindahkan ke kotak Frequency Variable (Gambar 8.6) dan klik OK.

(Catatan Jika data dicatat kasus per kasus dalam file data maka tidak akan ada kebutuhan untuk menggunakan prosedur Weight cases karena prosedur tab silang akan menghitung kasus secara otomatis.)



Gambar 8.6 Weight Cases dengan frekuensi dipindahkan pada Weight Cases by

Langkah 6 Pilih Analyze dari menu utama, klik Nonparametric dan kemudian Chi-square (Gambar 8.7). Akan muncul seperti pada (Gambar 8.8).

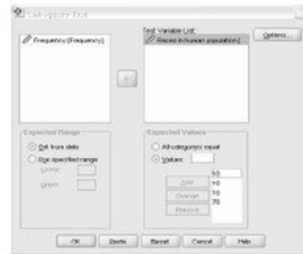


Gambar 8.7 Memilih Chi-Square dari menu utama



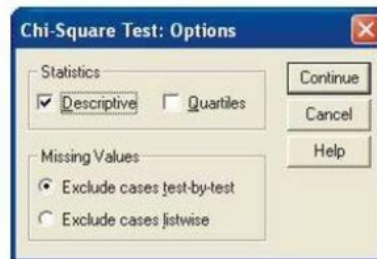
Gambar 8.8 Uji Chi-Square untuk memindahkan variabel

Langkah 7 Pindahkan "Ras dalam populasi manusia" dari sisi kiri ke **Test Variable List**. Kemudian pilih tombol Value pada Value yang diharapkan dan ketik rasio 10: 10:10:70 satu per satu dan tambahkan dengan mengklik Add sampai selesai seperti yang ditunjukkan pada Gambar 8.9. Sekarang pilih **Get from data** pada Expected Range dan periksa apakah semuanya sudah selesai seperti yang ditunjukkan pada Gambar 8.9. Klik OK untuk kembali ke Uji Chi-square dan klik Opsi.



Gambar 8.9 Uji Chi-Square dengan Value yang diharapkan

Langkah 8 Uji Chi-square: Opsi muncul seperti pada Gambar 8.10. Pilih Deskriptif dan kemudian klik Lanjutkan.



Gambar 8.10 Uji Chi-Square dengan pilihan Deskriptif pada Statistik

Langkah 9 Akhirnya klik OK untuk menjalankan analisis.

Output tampak seperti di bawah ini (**Output 1** dan **Output 2**)

Output 1 Frekuensi

Races in human population			
	Observed N	Expected N	Residual
Hispanic	24	20	4
Asians	11	20	-9
African American	20	20	0
White folk	145	140	5
Total	200		

Output 2

Test Statistics ^a	
	Races in human population
Chi-Square	5.03 ^a
df	3
Asymp. Sig.	0.17

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 20.0.

Interpretasi

Frekuensi yang diamati dan diharapkan diberikan dalam output 1. Nilai Chi-square adalah 5,029 untuk $df = 3$ dan signifikansi asimtotik adalah 0,170, (yaitu, p-value) diberikan dalam Output 2. Karena, p-value 0,170 lebih besar dari 0,05 ($p > 0,05$), perbedaan antara frekuensi yang diamati dan frekuensi yang diharapkan tidak signifikan. Oleh karena itu, Hipotesis nol (H_0), yaitu, tidak ada perbedaan antara frekuensi yang diamati dan yang diharapkan TIDAK DITOLAK pada tingkat signifikansi 5%. Jawaban atas masalahnya, yaitu bahwa komposisi ras pada sampel tidak berbeda secara signifikan dari nilai-nilai yang dihipotesiskan pada populasi di wilayah itu.

Contoh 8.2

Teori genetika memprediksi bahwa proporsi tanaman kacang di empat kelompok A, B, C dan D harus dalam rasio 9: 3: 3: 1. Jumlah tanaman dalam empat kelompok adalah $A = 365$, $B = 130$, $C = 125$ dan $D = 47$. Apakah hasil eksperimen ini mendukung teori bahwa hasilnya dalam rasio 9: 3: 3: 1?

Hipotesis Nol dan Alternatif

H_0 : Tidak ada perbedaan antara frekuensi yang diamati dan yang diharapkan.

H_A : Ada perbedaan nyata antara frekuensi yang diamati dan yang diharapkan.

Langkah 1 Klik Variabel view dan beri nama variabel di baris pertama sebagai "Pabrik", pilih numerik pada Type, pilih 0 di Desimal, beri label variabel pada kolom Label: "Karakter tanaman kacang". Di kolom Value, klik area abu-abu, untuk mendapatkan kotak dialog label Value dan tetapkan nomor Value

dan Label sebagai 1 untuk A, 2 untuk B, 3 untuk C, dan 4 untuk D dan klik OK.

Langkah 2 Masukkan Frekuensi di baris kedua di Name, pilih Numerik di Type, pilih 0 di Desimal dan beri label variabel di bawah kolom Label: Frekuensi.

Langkah 3 Klik Data View untuk beralih ke Data view dan masukkan data.

Langkah 4 Klik Data di menu utama, lalu klik Weight cases untuk membuka kotak Weight cases. Pilih Weight cases dengan tombol klik pada "Frekuensi" dan pindahkan ke **Frequency Variable** dan klik OK.

Langkah 5 Pilih Analyze dari menu utama kemudian klik Non-parametric dan kemudian Chi-square. Sebuah jendela muncul seperti pada Gambar 8.8. Pindahkan "Karakter dalam tanaman kacang" ke Test Variable List. Pilih Value pada Nilai yang diharapkan. Ketik rasio 9: 3: 3: 1, satu per satu dan tambahkan dengan mengklik saat sampai selesai. Sekarang pilih **Get from data** pada Expected range dan periksa apakah semuanya sudah selesai seperti yang ditunjukkan pada Gambar 8.11.



Gambar 8.11 Uji Chi-square dengan test variabel ditransfer, **Expected range** dipilih dan **Expected values**

Langkah 6 Klik OK. Output tampak seperti di bawah ini:

Output 1

Characters in pea plant			
	Observed N	Expected N	Residual
A-Round yellow	365	375.19	-10.19
B-Round green	130	125.06	4.94
C-Wrinkled yellow	125	125.06	0.06
D-Wrinkled green	47	41.69	5.31
Total	667		

Output 2

Test Statistics ^a	
	Characters in pea plant
Chi-Square	1.15 ^a
df	3
Asymp. Sig.	0.77

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 41.7.

Interpretasi

Output 1 memberikan Nilai yang diamati dan diharapkan. Output 2 memberikan hasil Chi-square.

Nilai untuk ¹3 df adalah 1,15. p-Value Asymp. Sig. adalah ²0,77. Karena p-value lebih besar dari 0,05 ($p > 0,05$), perbedaan antara frekuensi yang diamati dan yang diharapkan tidak signifikan dan oleh karena itu Hipotesis nol TIDAK DITOLAK. Hasil eksperimen mendukung teori dan hasilnya dalam rasio 9: ¹3: 3: 1.

"Goodness of Fit"distribusi Poisson dengan SPSS. Ada sejumlah kejadian seperti mendapatkan seorang anak perempuan atau anak laki-laki dalam satu kelahiran, jumlah penerimaan gawat darurat di rumah sakit pada selang waktu tertentu, jumlah kecelakaan di suatu tempat pada selang waktu tertentu, dll. Distribusi variabel ini dapat diberikan dalam bentuk probabilitas kejadian. Kita dapat menyesuaikan distribusi yang diberikan untuk memoduliskan distribusi seperti distribusi Poisson atau Binominal dan menguji "Goodness of Fit".

Sebelum memulai pada contoh, marilah kita terlebih dahulu memahami dasar-dasar, kondisi dan data yang berlaku pada hukum Poisson. Distribusi Poisson adalah distribusi teoritis diskrit. Dalam distribusi Poisson, kejadian n tidak diketahui, oleh karena itu seseorang dapat menghitung berapa kali suatu kejadian dapat terjadi tetapi bukan berapa kali kejadian gagal terjadi.

Suatu peristiwa dikatakan mengikuti distribusi Poisson ketika probabilitas kejadiannya sangat rendah, tetapi jumlah kejadiannya besar dan kejadiannya tidak bergantung pada ruang dan waktu.

Contoh peristiwa poisson sehubungan dengan waktu adalah jumlah kematian kecelakaan per tahun, jumlah gempa bumi per tahun, jumlah penerimaan gawat darurat dalam setahun dan seterusnya. Beberapa peristiwa poisson sehubungan dengan ruang adalah jumlah "tanaman lumut" di sisi bukit, jumlah organisme di ruang penghitungan, jumlah sel ragi di kotak dari haemocytometer dalam sampel dadih, jumlah penerimaan darurat di rumah sakit, jumlah kesalahan pencetakan dalam satu halaman, dll.

Dalam contoh-contoh ini kita dapat mengetahui apakah kejadian kejadian itu acak atau sebaliknya.

Contoh 8.3

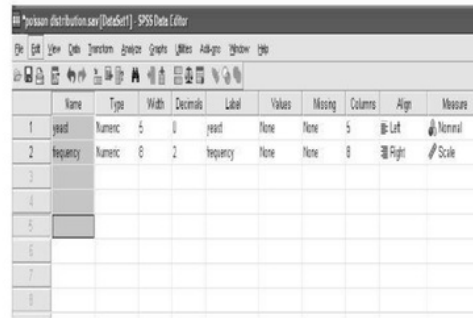
Tabel berikut memberikan distribusi sel ragi lebih dari 400 kotak haemocytometer dalam sampel dadih (curd). Hitunglah apakah distribusi ragi mengikuti hukum Poisson atau distribusi Poisson.

Yeast	0	1	2	3	4	5	6
Frequency	103	143	98	42	8	4	2

Kita perlu menghitung rata-rata, menemukan probabilitas Poisson dan mengujinya.

1. Menghitung rata-rata hitung distribusi.

Langkah 1 Klik Variabel view dari tab di kiri bawah. Ketik "Ragi" pada Name, pilih Numerik, pilih 0 di Desimal dan ketik di Label: "Jumlah ragi/persegi". Tidak perlu mengisi kolom Value. **31** da baris berikutnya pada Name : "frekuensi" dan isi sisanya seperti pada (Gambar 8.12).



Gambar 8.12 Variabel view dengan variabel diskrit dimasukkan

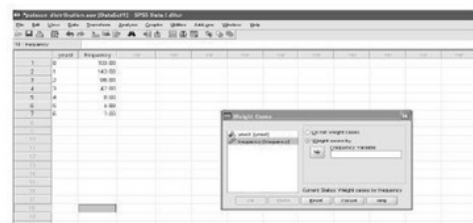
Langkah 2 Klik Data view dari tab di bagian bawah dan ketik data seperti yang diberikan pada Gambar 8.13.

	year	frequency
1	1	100.00
2	1	100.00
3	2	100.00
4	1	40.00
5	4	0.00
6	5	40.00
7	4	20.00
8		
9		

Gambar 8.13 Data view dengan frekuensi diskrit dimasukkan

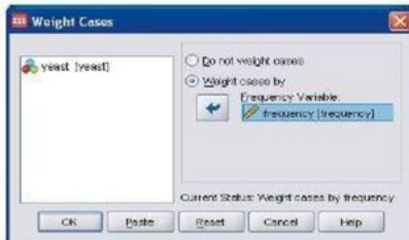
Karena variabel dikelompokkan, kita perlu menimbang kasus. Ragi adalah test variabel, bobot untuk setiap variabel adalah frekuensi. Oleh karena itu, berikan bobot sebagai berikut:

Langkah 3 lanjut ke Data dan pilih Weight cases, kotak dialog muncul seperti pada Gambar 8.14.



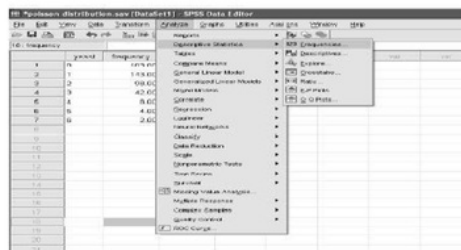
Gambar 8.14 Weight Cases dengan Data view

Sekarang klik tombol Weight Case di sebelah kiri. Pilih Frekuensi dari daftar dan klik tanda panah untuk mentransfernya ke kotak berlabel Variabel Frekuensi dan klik OK (Gambar 8.15).



Gambar 8.15 Weight Cases dengan frekuensi ditransfer pada Weight cases by

Langkah 4 Untuk menghitung rata-rata distribusi, klik Analisis, kemudian Statistik Deskriptif dan kemudian Frekuensi (Gambar 8.16).



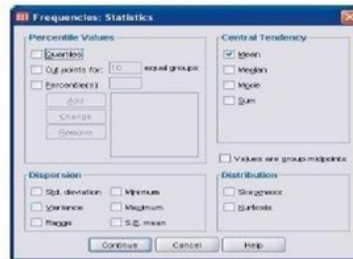
Gambar 8.16 Memilih frekuensi dari menu utama

Jendela muncul seperti di bawah ini. Pilih ragi dan klik pada panah untuk mentransfernya ke jendela Variabel (Gambar 8.17).



Gambar 8.17 Frekuensi dengan test variabel yang dipilih

Langkah 5 Klik Statistik dan pilih Mean pada Central Tendency (Kita perlu menghitung rata-rata hitung dari distribusi untuk menghitung distribusi probabilitas Poisson) (Gambar 8.18). Klik Lanjutkan dan klik OK. Hasilnya tampak seperti yang diberikan pada Output 1.



Gambar 8.18 Statistik dengan rata-rata hitung dipilih pada Tendensi sentral

Output 1

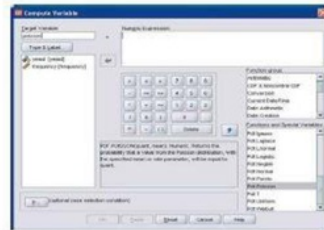
Statistics		
yeast		
N	Valid	400
	Missing	0
Mean		1.32

2. Perhitungan probabilitas Poisson

Langkah 6 Pilih **Transform** dari menu utama dan kemudian **Compute Variable** (Gambar 8.19).



Gambar 8.19 Memilih Compute Variable dari menu utama



Gambar 8.20 Compute Variable ketik Poisson dan pilih Pdf Poisson

Dalam Compute Variable, ketik nama Poisson pada **Target Variable**. Sekarang, giring ke bawah **Function group**: dan pilih PDF dan Noncentral PDF. Kemudian, gilir ke bawah Function dan **Special Variables**, dan pilih Pdf. Poisson. Pdf adalah singkatan dari "Probability density function" (Gambar 9.20). Klik pada panah ke atas untuk memindahkannya ke jendela kanan atas bernama **Numeric Expression** (Gambar 8.21).



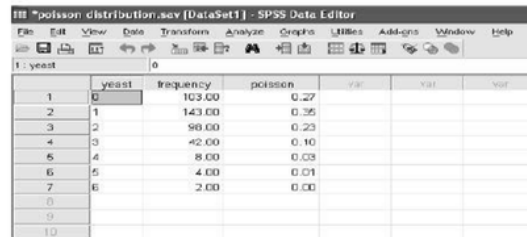
Gambar 8.21 Compute Variable untuk mentransfer Variabel dan ketik nilai Rata-rata hitung

Sekarang jendela Ekspresi Numerik memiliki Pdf. Poisson. Pilih lagi dari jendela kiri dan klik panah untuk mentransfernya ke jendela kanan. Ketik nilai rata-rata hitung (1,3225) dari Output 1 (Gambar 8.22). Periksa prosedur ini dengan benar, karena ini adalah dasar untuk perhitungan probabilitas Poisson. Terakhir klik OK.



Gambar 8.22 Compute Variable dengan semua opsi selesai

Output tampak seperti diberikan di bawah ini dengan probabilitas Poisson dalam tampilan data SPSS bersama dengan data yang sudah dimasukkan (Gambar 8.23). Sekarang, hapus seluruh baris (baris ke-7) di mana nilai poisson adalah 0. (Dalam masalah yang melibatkan hukum Poisson, jika Value Poisson adalah '0' maka kita tidak dapat memasukkan Value itu.)



yeast	frequency	poisson
1	103.00	0.27
2	143.00	0.35
3	98.00	0.23
4	42.00	0.10
5	8.00	0.03
6	4.00	0.01
7	2.00	0.00
8		
9		
10		

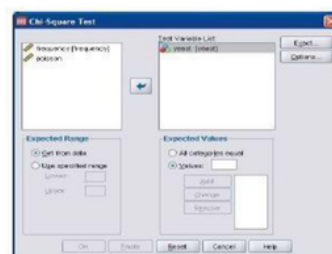
Gambar 8.23 Tampilan Data dengan probabilitas Poisson yang dihitung dari opsi Transform

3. Analisis Kesesuaian dari Analysis of Goodness of fit

Hipotesis nol: Tidak ada perbedaan antara frekuensi teoritis Poisson yang diamati dan yang diharapkan.

Hipotesis alternatif: Ada perbedaan antara frekuensi teoritis Poisson yang diamati dan yang diharapkan.

Langkah 7 Pilih Analisis dari menu utama, lalu klik Nonparametric dan kemudian Chi-square. Akan muncul seperti pada Gambar 8.24, pilih ragi dan transfer ke Test Variable List.



Gambar 8.24 Kotak Uji Chi-Square

Langkah 8 Pilih **Get from Data** di tombol Expected Range. Pilih tombol radio **Value** pada Expected Value.

Masukkan probabilitas Poisson dari Data Editor ke kotak Value dan tambahkan satu per satu sampai selesai seperti pada Gambar 8.25.



Gambar 8.25 Kotak Uji Chi-Square dengan Test Variable dan Expected Value (probabilitas Poisson)

Langkah 9 Klik OK. Output muncul di jendela output.

Output 2 Frekuensi

yeast			
	Observed N	Expected N	Residual
0	103	108.55	-5.55
1	143	140.71	2.29
2	98	92.46	5.54
3	42	40.2	1.8
4	8	12.06	-4.06
5	4	4.02	-0.02
Total	398		

Output 3

Test Statistics	
	yeast
Chi-Square	2.1 ^a
df	5
Asymp. Sig.	0.84

a. 1 cells (16.7%) have expected frequencies less than 5. The minimum expected cell frequency is 4.0.

Interpretasi

Output 1 memberikan frekuensi yang diamati dan diharapkan.

Nilai Chi-square adalah 2,100 untuk df 5 dan signifikansi asimtotik adalah 0,835 (Output 2), yang lebih besar dari 0,05 ($p > 0,05$). Tidak ada perbedaan yang signifikan antara distribusi yang diberikan dan distribusi probabilitas teoretis yang diharapkan. Oleh karena itu, Hipotesis nol TIDAK DITOLAK dan distribusi yang diberikan mengikuti hukum Poisson. Terdapat kesesuaian yang pas. Secara ilmiah, probabilitas mendapatkan sel ragi dalam kamar manapun dari ruang penghitungan sangat rendah tetapi jumlah total sel ragi dalam sampel sangat tinggi.

UJI CHI-SQUARE UNTUK INDEPENDENSI ATRIBUT DENGAN SPSS

Uji Chi-square untuk independensi adalah uji untuk mencari tahu apakah dua variabel kategori terkait satu sama lain. Untuk melakukan Uji Chi-square seperti ini kita memerlukan dua variabel kategori yang dapat secara wajar diberikan dalam tabel bivariat (yaitu, dengan sejumlah kategori terbatas).

Contoh 8.4

Sekelompok mahasiswa diklasifikasikan dalam hal jenis kelamin (pria dan wanita) dan golongan darah (A, B, AB dan O). Buktikan apakah ada hubungan antara gender dan golongan darah.

Sex	Blood group				Total
	A	B	AB	O	
Male	28	16	7	49	100
Female	25	13	8	54	100
Total	53	29	15	103	200

Dua pengelompokan variabel kategori

- i. Gender – pria dan wanita
- ii. Golongan darah – A, B, AB dan O.

Hipotesis

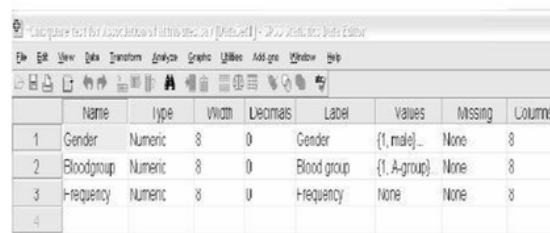
H_0 : Tidak ada hubungan antara gender dan golongan darah.

H_A : Ada hubungan antara gender dan golongan darah.

Langkah 1 Buka file data baru di SPSS.

Langkah 2 Beri nama variabel dalam Variabel view. Klik Variabel view, ketik variabel di baris pertama "Gender" pada Name, pilih Numerik di Type, pilih 0 di Desimal, label variabel di kolom Label. Di kolom Value, klik area abu-abu, untuk mendapatkan kotak dialog label Value dan tetapkan nomor dan label Value 1 untuk pria dan 2 untuk wanita. Di baris kedua, ketik variabel kedua sebagai "Blood Group" di kolom Name, pilih Numerik di bawah Type, pilih 0 di Desimal, beri label variabel "Blood Group" di kolom Label. Di kolom Value, klik area abu-abu untuk mendapatkan kotak dialog label Value dan tetapkan nomor Value dan beri label 1 untuk A, 2 untuk B, 3 untuk AB dan 4 untuk O (Gambar 8.25).

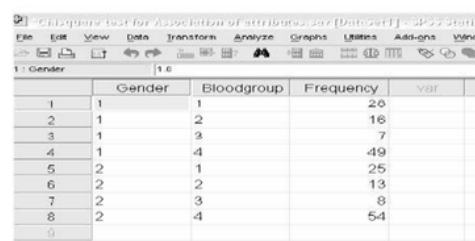
Langkah 3 Pada baris ketiga Frekuensi pada Name, pilih Numerik di bawah Type, pilih 0 di kolom Desimal, kolom label sebagai "Frekuensi". Tidak perlu memberikan label Value di Value (Gambar 8.26).



	Name	Type	Width	Decimals	Label	Values	Missing	Columns
1	Gender	Numeric	8	0	Gender	(1,male)...	None	8
2	Bloodgroup	Numeric	8	0	Blood group	(1,A-group)	None	8
3	Frequency	Numeric	8	0	Frequency	None	None	8
4								

Gambar 8.26 Variabel view dengan nama variabel

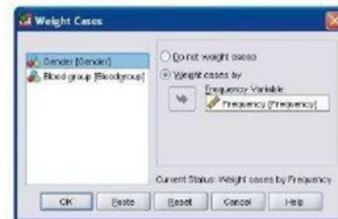
Langkah 4 Klik pada Data view dan masukkan data di bawah masing-masing variabel yang berlabel Variabel view seperti pada Gambar 8.27.



	Gender	Bloodgroup	Frequency	var
1	1	1	20	
2	1	2	16	
3	1	3	7	
4	1	4	49	
5	2	1	25	
6	2	2	13	
7	2	3	8	
8	2	4	54	
9				

Gambar 8.27 Data view dengan frekuensi yang dimasukkan untuk dua variabel kategorial

Langkah 5 Pilih Data dari menu utama dan kemudian klik Weight case. **1** i akan membuka kotak dialog Weight case (Gambar 8.28), pilih Weight cases by dan transfer Frequency tke Frequency Variable. Klik OK. Sekarang seluruh tampilan akan hilang.



Gambar 8.28 Kotak dialog Weight Cases

Catatan: Seperti dinyatakan sebelumnya jika data telah dicatat kasus per kasus dalam file data maka tidak akan ada kebutuhan untuk menggunakan prosedur weight case karena prosedur tab silang akan menghitung kasus secara otomatis.

Langkah 6 Pilih Analisis dari menu utama lalu klik Statistik deskriptif lalu klik Crosstabs. Crosstabs muncul seperti pada Gambar 8.29. Transfer Gender dan Kelompok Darah di baris dan kotak kolom masing-masing. Jika Anda ingin diagram batang pilih Display cluster bar chart. Klik tombol Statistik.



Gambar 8.29 Crosstabs untuk memilih variabel kategori dalam Baris dan Kolom

Langkah 7 Crosstabs: Jendela statistik muncul seperti pada Gambar 8.30. Pilih opsi Chi-square dan lainnya sesuai kebutuhan. Klik Lanjutkan.



Gambar 8.30 Crosstabs: Statistik dengan opsi Chi-Square

Langkah 8 Terakhir, klik OK untuk analisis. Output SPSS muncul seperti di bawah ini dan tafsirkan hasilnya.

Output 1

Gender * Blood group Crosstabulation

Count		Blood group				
		A-group	B-group	AB-group	O-group	Total
Gender	male	28	16	7	49	100
	female	25	13	8	54	100
Total		53	29	15	103	200

Output 2

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	0.79 ^a	3	0.85
Likelihood Ratio	0.79	3	0.85
Linear by Linear Association	0.58	1	0.45
N of Valid Cases	200		

a. 0 cells (.0%) have expected count less than 5.
The minimum expected count is 7.50.

Interpretasi

Output 1 memberikan frekuensi kelompok darah pada pria dan wanita dalam bentuk tabel (tabulasi silang) untuk $df = 3$ yaitu, $[(r - 1) (c - 1)] = [(2 - 1) (4 - 1)] = 3$, p-value 0,852 lebih besar dari 0,05. Perbedaan dianggap tidak signifikan. Hipotesis nol TIDAK DITOLAK dan oleh karena itu, tidak ada hubungan antara jenis dan golongan darah. Dengan kata lain, jenis kelamin dan golongan darah bersifat independen pada manusia.

Output 2 memberikan hasil Uji Chi-Square.

Contoh 8.5

Dua kelompok A dan B terdiri dari 100 orang masing-masing memiliki penyakit tertentu. Obat diberikan pada kelompok A dan bukan obat untuk kelompok B. Jika tidak keduanya diperlakukan identik. Ditemukan bahwa pada kelompok A dan B, 80 dan 65 orang telah sembuh dari penyakit. Ujilah Hipotesis bahwa obat membantu menyembuhkan penyakit pada tingkat signifikansi 0,05.

Hipotesis

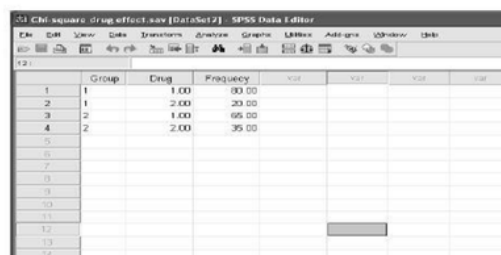
H0: Obat tidak efektif dalam menyembuhkan penyakit.

HA: Obat efektif dalam menyembuhkan penyakit.

Langkah 1 Buka file data baru di SPSS.

Langkah 2 Beri nama variabel dalam Variabel view. Klik Variabel view, beri nama variabel di baris pertama: "Grup" pada Name, pilih Numerik di pada Type, pilih 0 di Desimal, beri label variabel di kolom Label. Di kolom Value, klik area abu-abu, untuk mendapatkan kotak dialog label Value dan tetapkan nomor dan label Value, 1 untuk grup A dan 2 untuk grup B. Beri nama variabel kedua: "Drug" di baris kedua di kolom Name, pilih Numeric di Type, pilih 0 di Desimal, beri label variabel di bawah kolom Label. Di kolom Value, klik area abu-abu untuk mendapatkan kotak dialog label Value dan tetapkan nomor Value dan label sebagai 1 " for cured" (untuk disembuhkan) dan 2 "not for cured" (untuk tidak disembuhkan) (Gambar 9.31).

Langkah 3 Pada baris ketiga, ketik "Frekuensi" pada Name, pilih Numerik pada Type, pilih 0 di Desimal, kolom Label sebagai "Frekuensi". Tidak perlu memberikan label Value di bawah Value.



	Group	Drug	Frequency				
1	1	1.00	80.00				
2	1	2.00	20.00				
3	2	1.00	65.00				
4	2	2.00	35.00				
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							

Gambar 8.31 Data view dengan frekuensi untuk dua variabel kategori

Langkah 4 dan 5 Merujuk contoh sebelumnya dan menyelesaikan weight case.

Langkah 6 Pilih Analisis dari menu utama lalu klik Statistik deskriptif lalu klik Crosstabs. Dalam Crosstabs, pindahkan Group dan Drug dalam baris dan kotak kolom masing-masing. Jika Anda ingin diagram bar pilih Display cluster bar chart dan klik tombol Statistik.

Langkah 7 Crosstabs: Jendela statistik muncul, pilih Chi-square dan opsi lainnya sesuai kebutuhan. Klik Lanjutkan dan OK. Output tampak seperti di bawah ini.

Output 1

Group * Drug Crosstabulation

Count		Drug		Total
		Cured	not cured	
Group	Group A	80	20	100
	Group B	65	35	100
Total		145	55	200

Output 2

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.643 ^a	1	.018		
Continuity Correction ^b	4.915	1	.027		
Likelihood Ratio	5.698	1	.017		
Fisher's Exact Test				.026	.013
Linear-by-Linear Association	5.614	1	.018		
N of Valid Cases	200				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 27.50.
b. Computed only for a 2x2 table.

Interpretasi

Pearson untuk $df = 1$, yaitu, $[(r - 1) (c - 1)] = [(2 - 1) (2 - 1)]$. Value p 0,018 kurang dari 0,05. Oleh karena itu, Hipotesis nol ditolak dan obat ini efektif dalam menyembuhkan penyakit ($p < 0,05$).

REVIEW LATIHAN

1. Sesuaikan distribusi Poisson untuk data berikut pada jumlah plankton dari sampel air dan Uji Goodness of fit.

Number of plankton	0	1	2	3	4	5	6
Number of samples	93	105	89	42	11	8	2

2. Data berikut memberikan respon obat tertentu dalam menyembuhkan penyakit. Selidikilah apakah obat itu efektif dalam menyembuhkan penyakit.

	Cured	Not cured
Drug	300	200
No drug	50	250

4. Data berikut memberikan warna mata ayah dan warna mata anak laki-laki. Apakah ada hubungan antara warna mata anak laki-laki dengan ayah mereka.

Eye colour of father	Eye colour of son	
	Dark	Light
Dark	230	148
Light	150	471

5. Mahasiswa dari perguruan tinggi dinilai berdasarkan IQ dan kondisi ekonomi mereka. Periksalah apakah ada hubungan antara IQ dan kondisi ekonomi.

Economic conditions	IQ	
	High	
Rich	460	140
Poor	240	160

REFERENSI

37

Andy Field. *Discovering Statistics Using SPSS*, 2nd edition. Sage Publications, London, 2005.

39

Daniel, Wayne B. *Biostatistics: A Foundation for Analysis in the Health Sciences*, 2nd Edition. John Wiley & Sons, New York, 1983.

1

Eelko Huizingh. *Applied Statistics with SPSS*. Sage Publications, New Delhi, India, 2007.

1

Ipsen, Johannes and Feigl, Polly. *Bancroft's Introduction to Biostatistics*, Harper International Edition. Harper and Row Publishers, New York, 1970.

Jacobson, Perry E. *Introduction to Statistical Measure for Social and Behavioural Sciences*. Dryden Press, New York, 1976.

Jerrold, H. Zar. *Biostatistical Analysis*, 4th edition. Prentice Hall International, New Jersey, 1999.

John, E. Freund. *Statistics: A First Course*, 3rd edition. Prentice–Hall, Inc. Englewood Cliffs, New Jersey, 1981.

Mandal, R.B. *Statistics for Geographers and Social Scientists*. Concept Publishing Company, New Delhi, 2001

Paul, R. Kinnear and Colin, D. Gray. *SPSS 14 Made Simple*. Psychology Press, Taylor and Francis Group, Hove and New York, 2006.

Rajathi, A and Chandran P. *SPSS for You*. MJP Publishers Tamilnadu, India, 2015

1

Sokal, R. Robert, and Rohlf, F. James. *Introduction to Biostatistics*. W.H. Freeman and Company, San Francisco, 1973.

GLOSARIUM

Adjusted R^2 Ukuran kehilangan kekuatan prediktif dalam analisis regresi. R disesuaikan yang memberitahu kita berapa banyak varians hasil yang akan dipertanggungjawabkan, jika model telah diturunkan dari populasi dari mana sampel diambil.

Alternative hypothesis Setiap pernyataan (hipotesis) yang melengkapi hipotesis nol. Ini menyatakan bahwa rata-rata sampel dan rata-rata populasi tidak sama.

ANOVA. adalah prosedur statistik inferensial yang meneliti perbedaan atau perbedaan antara kelompok kontrol dan kelompok perlakuan dalam penyelidikan.

Bivariate correlation Korelasi (hubungan) antara dua variabel.

Categorical variable Setiap variabel yang terdiri dari kategori obyek atau entitas. Hasil tes dalam kelas adalah contoh yang baik karena diklasifikasikan ke dalam sukses dan gagal.

Chi-square distribution Distribusi probabilitas jumlah kuadrat dari beberapa variabel terdistribusi normal. Ini digunakan untuk menguji hipotesis tentang variabel kategori.

Chi-square test Secara umum mengacu pada uji Chi-square Pearson. Digunakan untuk mencari perbedaan antara frekuensi yang diamati dan yang diharapkan berdasarkan beberapa model atau untuk menguji independensi dari dua variabel kategori.

Confidence interval Rangkaian nilai di sekitar parameter statistik tersebut (misalnya, rata-rata) yang diyakini mengandung probabilitas tertentu (95%), nilai sebenarnya dari parameter statistik tersebut (mis. Rata-rata populasi).

Contingency table Sebuah tabel yang mengklasifikasikan individu sehubungan dengan dua atau lebih variabel kategori. Tingkat masing-masing variabel diatur dalam baris dan kolom dan jumlah individu yang masuk ke dalam setiap kategori dicatat dalam sel tabel. Misalnya, jika mahasiswa di sebuah perguruan tinggi diklasifikasikan berdasarkan jenis kelamin dan golongan darah, tabel kontingensi akan menunjukkan jumlah laki-laki dalam golongan darah A, jumlah perempuan dalam golongan darah B dan sebagainya.

Correlation coefficient Angka desimal antara 0 dan 1 (negatif atau positif), yang menunjukkan tingkat dan arah hubungan dua variabel kuantitatif dan diwakili oleh r .

Covariance Ukuran seberapa besar penyimpangan dari dua variabel yang cocok.

Covariate Sebuah variabel yang terkait dengan variabel hasil yang diukur. Pada dasarnya, apa pun yang berdampak pada variabel dependen dan yang tidak dapat dikontrol oleh desain dapat menjadi kovariat.

Criterion variable Variabel hasil (dependen) yang diprediksi dalam analisis regresi atau penelitian korelasi.

Data Editor Jendela utama SPSS untuk memberi nama variabel, memasukkan data, dan melakukan analisis.

Data View Salah satu dari dua cara untuk melihat konten data editor. Tampilan data memiliki lembar kerja/spreadsheet untuk memasukkan data.

Degrees of Freedom Suatu hal yang sulit untuk didefinisikan dalam glosarium. Ini adalah jumlah item yang bebas bervariasi ketika memperkirakan beberapa parameter statistik. Derajat kebebasan (df) diberikan oleh jumlah pengamatan independen dikurangi jumlah parameter yang diperkirakan. Jika ada n pengamatan dalam sampel dan penyimpangan semua n item diperkirakan dari rata-rata sampel dan hanya satu rata-rata parameter diperkirakan maka $df = n - 1$. Ini adalah istilah yang dipinjam dari ilmu fisika, di mana tingkat kebebasan suatu sistem adalah jumlah kendala yang diperlukan untuk menentukan keadaan sepenuhnya pada titik mana pun. atau disebut sebagai variabel hasil. Istilah ini digunakan dalam penelitian apa pun untuk menentukan hubungan sebab-akibat.

Dependent variable Variabel yang diukur dalam penelitian berkenaan dengan variabel yang dimanipulasi oleh yang melakukan eksperimen (experimenter). Juga disebut sebagai variabel hasil (tidak bebas). Istilah ini digunakan dalam penelitian apa pun untuk menentukan hubungan sebab dan akibat.

Descriptive Statistics Prosedur statistik yang hanya menggambarkan berbagai karakteristik data daripada mencoba menyimpulkan sesuatu dari data.

Factor Nama lain untuk variabel independen (bebas) atau prediktor yang digunakan dalam menggambarkan desain eksperimental.

F-Ratio Suatu uji statistik dengan distribusi probabilitas yang diketahui (Distribusi-F). Ini adalah rasio variabilitas rata-rata dalam data yang dapat dijelaskan oleh model tertentu terhadap variabilitas rata-rata yang tidak dijelaskan oleh model yang sama.

Goodness of fit Suatu indeks untuk menemukan seberapa baik suatu model sesuai dengan data dari mana ia dihasilkan. Tes chi-square adalah salah satu tes untuk menemukan kesesuaian yang paling pas.

Grouping variable Jika pengamatan dalam suatu penelitian dikumpulkan sesuai dengan kesamaan atau kemiripan, masing-masing membentuk kelompok. Dalam SPSS, satu set nomor kode diberikan untuk menunjukkan setiap kelompok dalam tampilan variabel.

Hypothesis testing Suatu prosedur statistik untuk menguji hipotesis nol terhadap hipotesis alternatif. Prosedur statistik dalam pengujian hipotesis meliputi t, F dan uji X^2 (Chi-Square)

Hypothesis Dalam statistik, hipotesis adalah pernyataan tentang suatu populasi, seperti sifat distribusi. (Ada dua jenis hipotesis: Hipotesis Nol H_0 dan Hipotesis Alternatif - H_A).

Independent Variable Suatu variabel yang diteliti, untuk menentukan apakah ia memiliki efek kausal pada variabel dependen. Dalam studi regresi, istilah ini digunakan untuk menunjukkan variabel prediktor atau regresi.

Inferential statistics Termasuk seperangkat alat statistik yang memungkinkan peneliti untuk menyimpulkan atau membuat kesimpulan tentang data.

Kendalls tau Uji korelasi digunakan dengan dua variabel ordinal atau variabel ordinal dan interval. Sebelum ke komputer, rho lebih disukai daripada tau karena kemudahan komputasi. Sekarang komputer telah membuat perhitungan menjadi mudah sehingga tau umumnya lebih disukai. Tau Kendall secara parsial juga tersedia sebagai analog ordinal dengan korelasi Pearson parsial.

MANOVA Analisis varian multivariat. Ini adalah analisis varians (ANOVA) yang berlaku untuk multivariat (kumpulan data

multivariat yang berisi pengamatan pada situasi tiga atau lebih variabel (variabel dependen), untuk dua atau lebih variabel independen.

Mean Rata-rata dari sekumpulan angka atau skor dalam suatu distribusi. Untuk mendapatkan nilai rata-rata, semua nilai ditambahkan dan jumlahnya dibagi dengan jumlah total semua nilai.

Measures of central tendency Angka tunggal yang digunakan untuk menggambarkan kumpulan data yang lebih besar dalam skor distribusi. Ukuran tendensi sentral adalah mean, median, dan modus.

Median Skor atau angka, yang jatuh langsung di tengah distribusi angka atau membagi data menjadi dua bagian yang sama.

Mode Modus, jumlah skor yang paling sering terjadi dalam distribusi angka.

Multiple regression Suatu perpanjangan dari regresi sederhana di mana suatu hasil diprediksi oleh kombinasi linear dari dua atau lebih variabel predictor (bebas).

Multivariate analysis Analisis varians yang melibatkan lebih dari satu variabel hasil yang telah diukur.

Multivariate Analisis varians yang melibatkan lebih dari satu variabel hasil (multivariat = banyak variabel)

Negatively Skewed adalah asimetri dalam suatu distribusi. Distribusi yang condong negatif memiliki sebagian besar skornya berkumpul di ujung yang lebih tinggi (sisi kanan) dari distribusi.

Nominal data Suatu set data yang angka-angkanya hanya mewakili nama, angka-angka tidak memiliki arti selain nama.

Non-parametric tests Uji statistik atau prosedur yang tidak mengasumsikan bahwa data berasal dari distribusi normal. Uji seperti tes Mann-Whitney dan χ^2 -test adalah uji non-parametrik.

Normal distribution Nama lain untuk kurva berbentuk lonceng; distribusi ini membentuk konsep yang sangat penting dalam statistik dan penelitian. Ini memiliki karakteristik tertentu

yang berbeda, yang menjadikannya alat yang sangat berguna untuk statistik deskriptif dan inferensial.

Null hypothesis Hipotesis Nol, sebuah hipotesis yang dibuat oleh peneliti tentang masalah yang sedang diselidiki. Hipotesis nol adalah pernyataan 'tidak ada perbedaan'. H_0 menyatakan bahwa tidak ada perbedaan yang signifikan antara rata-rata sampel dan rata-rata populasi, atau antara rata-rata dua populasi, atau antara rata-rata lebih dari dua populasi.

One-tailed test Uji satu-ekor digunakan untuk menguji apakah rata-rata sampel secara signifikan lebih besar dari rata-rata populasi atau jika secara signifikan kurang dari itu, tetapi tidak keduanya.

p-value Nilai probabilitas (p-value) dari suatu uji hipotesis statistik. Ini adalah probabilitas mendapatkan nilai statistik uji yang ekstrem seperti yang diamati secara kebetulan. Nilai p yang kecil menunjukkan bahwa hipotesis nol ditolak. Semakin kecil nilai-p, semakin meyakinkan penolakan terhadap hipotesis nol. p-value menunjukkan kekuatan bukti untuk menolak hipotesis nol H_0 .

Percentile Distribusi apa pun dapat dijelaskan dalam persentil. 10th persentil adalah nilai dalam distribusi di bawah ini yang terletak 10% dari nilai. Persentil ke-90 adalah nilai di bawah 90% dari nilai-nilai tersebut. Jadi 50 persen adalah median dalam suatu distribusi.

Positively skewed Kecondongan yang positif adalah asimetri dalam suatu distribusi. Distribusi yang condong positif memiliki sebagian besar skornya berkumpul di ujung bawah (kiri) distribusi.

Post-hoc comparisons Perbandingan yang tidak terencana, seseorang ingin membuat setelah data telah dikumpulkan. Biasanya dilakukan jika ada perbedaan yang signifikan antara dua kelompok (perbandingan berpasangan) yang meliputi berbagai tingkat variabel yang diteliti.

Predictor variable Variabel dari mana variabel kriteria ditemukan dalam studi prediksi.

Probability Probabilitas istilah tidak memungkinkan definisi konkret. Dapat didefinisikan dengan berbagai cara. Secara sederhana, probabilitas adalah rasio jumlah hasil yang

menguntungkan dibagi dengan jumlah hasil dalam satu percobaan atau eksperimen.

Qualitative variable Karakter atau properti, seperti golongan darah, jenis kelamin atau sukubangsa, yang dapat dinyatakan dalam bentuk barang dan bukan barang dalam jumlah. Ini adalah atribut dan bersifat deskriptif.

Quantitative variable: Variabel yang statusnya berbeda dapat dinyatakan dalam angka. Karakteristik seperti tinggi, berat, panjang, dll. Diukur secara kuantitatif.

Regression Proses matematis menggunakan pengamatan untuk memprediksi variabel target-dependen (kriteria) dari variabel lain yang dikenal sebagai regressor atau variabel independen. Prediksi ini dibuat dengan membangun persamaan regresi atau garis regresi, garis yang paling sesuai melalui data.

Sample: Kumpulan unit yang lebih kecil dari populasi yang digunakan untuk menentukan karakteristik atau kebenaran populasi.

Sampling distribution Distribusi probabilitas dari suatu parameter statistik. Jika sejumlah besar sampel diambil dari populasi dan beberapa statistik dihitung, misalnya rata-rata, kita akan membuat distribusi frekuensi rata-rata. Distribusi yang dihasilkan akan membentuk distribusi sampling nilai rata-rata.

Scatterplot adalah diagram titik/pencar pada sumbu koordinat (sumbu X dan Y) yang digunakan untuk mewakili dan menggambarkan hubungan antara dua variabel kuantitatif.

Skewed Sebuah distribusi miring jika mayoritas skor dikelompokkan di satu ujung atau yang lain.

Spearman's rho Korelasi paling umum untuk digunakan dengan dua variabel ordinal atau variabel ordinal dan interval. Rho untuk data peringkat sama dengan r Pearson untuk data peringkat.

Standard deviation Perkiraan sebaran rata-rata (variabilitas) dari sekumpulan data yang diukur dalam satuan pengukuran yang sama dengan data asli. Ini adalah akar kuadrat dari varians.

Standard error Deviasi standar dari distribusi sampling dari suatu statistik. Sebagai contoh, rata-rata adalah parameter statistik, standar error menjelaskan seberapa besar rata-rata

sampel berbeda dari rata-rata populasi. Semakin besar standar error, semakin besar kemungkinan bahwa sampel mungkin bukan merupakan cerminan akurat dari populasi dari mana sampel berasal.

String variable Variabel yang melibatkan kata, mis., String huruf (mis., Jenis kelamin, golongan darah, dll.).

Sum of squares Perkiraan variabilitas total dari satu set data. Pertama penyimpangan untuk setiap skor dihitung dan kemudian nilai ini dikuadratkan dan disimpulkan. Ini dilambangkan dengan SS.

Syntax Perintah tertulis yang ditentukan sebelumnya yang menginstruksikan SPSS apa yang ingin dilakukan pengguna.

Total sum of squares Perkiraan variabilitas total dalam satu set data. Ini adalah penyimpangan kuadrat total antara setiap pengamatan dan dari rata-rata keseluruhan dari semua pengamatan.

t-test Prosedur statistik untuk menemukan signifikansi perbedaan antara dua kelompok. Nilai rata-rata kedua kelompok dibandingkan satu sama lain.

Two-tailed test Uji dua sisi rata-rata bahwa tingkat signifikansi 0,05, terdistribusi secara merata di kedua sisi ekor. 0,025 ada di setiap ekor distribusi uji statistik. Uji dua sisi akan menguji apakah rata-rata secara signifikan lebih besar dari μ atau jika rata-rata secara signifikan kurang dari μ .

Type I error Menolak hipotesis nol ketika itu sebenarnya benar. Probabilitas Tipe I, kesalahan adalah tingkat signifikansi dan juga dikenal sebagai alfa.

Type II error Tidak menolak hipotesis nol ketika itu sebenarnya salah. Probabilitas kesalahan Tipe II alfa dikenal sebagai tingkat beta atau tingkat beta. Tingkat beta ditentukan oleh sejumlah faktor seperti ukuran sampel dan tingkat signifikansi.

Univariate Rata-rata hanya satu variabel. Istilah ini biasanya merujuk pada situasi di mana hanya satu variabel yang diukur.

Variable view Salah satu dari dua cara untuk melihat konten data editor. Ini memiliki spreadsheet untuk memasukkan nama dan detail variabel.

Variable Sebuah properti atau karakteristik tempat pengumpulan data. Ada variabel kualitatif dan kuantitatif.

Variance Jumlah rata-rata dispersi atau penyebaran dalam distribusi. Penyimpangan dari rata-rata kuadrat dan disimpulkan untuk menemukan varians.

Wilcoxon signed-rank test Sebuah ¹⁵ uji non-parametrik yang digunakan untuk menguji perbedaan antara dua sampel terkait. Ini adalah non-parametrik yang setara dengan uji-t berpasangan

Z-score Sebuah konversi skor mentah menjadi skor terstandarisasi yang diwakili dalam satuan standar deviasi. Ini adalah prosedur statistik yang umum digunakan yang digunakan untuk membandingkan beberapa uji yang mungkin tidak diukur pada skala yang sama.

PENGOLAHAN DAN ANALISIS DATA DENGAN SPSS

ORIGINALITY REPORT

6%

SIMILARITY INDEX

6%

INTERNET SOURCES

1%

PUBLICATIONS

1%

STUDENT PAPERS

PRIMARY SOURCES

1

www.fb4all.com

Internet Source

2%

2

www.scribd.com

Internet Source

1%

3

dlds.jugem.jp

Internet Source

<1%

4

harmoko-fakultaspertanianunisri.blogspot.com

Internet Source

<1%

5

eprints.uny.ac.id

Internet Source

<1%

6

docplayer.info

Internet Source

<1%

7

pt.scribd.com

Internet Source

<1%

8

media.neliti.com

Internet Source

<1%

9

repository.usu.ac.id

Internet Source

<1%

10	Submitted to Fakultas Ekonomi Universitas Indonesia Student Paper	<1 %
----	--	------

11	edoc.site Internet Source	<1 %
----	------------------------------	------

12	primadonakita.blogspot.com Internet Source	<1 %
----	---	------

13	es.scribd.com Internet Source	<1 %
----	----------------------------------	------

14	Submitted to Universitas Jember Student Paper	<1 %
----	--	------

15	repository.uinjkt.ac.id Internet Source	<1 %
----	--	------

16	www.slideshare.net Internet Source	<1 %
----	---------------------------------------	------

17	eprints.undip.ac.id Internet Source	<1 %
----	--	------

18	Submitted to Universitas Pendidikan Indonesia Student Paper	<1 %
----	--	------

19	siregarbakti.blogspot.com Internet Source	<1 %
----	--	------

20	sir.stikom.edu Internet Source	<1 %
----	-----------------------------------	------

21

Internet Source

<1 %

22

fr.scribd.com

Internet Source

<1 %

23

id.123dok.com

Internet Source

<1 %

24

etheses.uin-malang.ac.id

Internet Source

<1 %

25

repository.upi.edu

Internet Source

<1 %

26

weblogs.hianoto.net

Internet Source

<1 %

27

taripanya.blogspot.com

Internet Source

<1 %

28

Salim A., Colfer C.J.P., McDougall C.. "Panduan cara pemberian skor dan analisis untuk menilai kesejahteraan manusia", Center for International Forestry Research (CIFOR), 2000

Publication

<1 %

29

dokumen.tips

Internet Source

<1 %

30

www.greenilmu.com

Internet Source

<1 %

31

documents.mx

Internet Source

<1 %

32

vdocuments.mx

Internet Source

<1 %

33

vdocuments.site

Internet Source

<1 %

34

ameliafajars.blogspot.com

Internet Source

<1 %

35

Submitted to Universitas Muria Kudus

Student Paper

<1 %

36

asemanis.com

Internet Source

<1 %

37

Pedro Gabriel Ferreira, Paulo J. Azevedo.
"Evaluating Protein Motif Significance
Measures: A Case Study on Prosite Patterns",
2007 IEEE Symposium on Computational
Intelligence and Data Mining, 2007

Publication

<1 %

38

Submitted to Universitas Putera Batam

Student Paper

<1 %

39

www.physics.nus.edu.sg

Internet Source

<1 %

40

widyo.staff.gunadarma.ac.id

Internet Source

<1 %

41

eau.sagepub.com

Internet Source

<1 %

42

Chendi Liana, Herta Armianti, Muhammad Ali.
"Pengaruh Terpaan Tayangan Religi di Televisi
Terhadap Sikap dan Perilaku Taqwa Pemirsa",
Jurnal Dakwah dan Komunikasi, 2017

Publication

<1 %

43

id.scribd.com

Internet Source

<1 %

44

mafiadoc.com

Internet Source

<1 %

45

Tulus Bangun Hutagalung. "Sustainability
Implication of the Lestari Tea Certification
Program Implementation on the Tea Small
Holder at Campaka Subdistrict Cianjur", Jurnal
Penelitian Teh dan Kina, 2016

Publication

<1 %

46

biologi.fkip.uns.ac.id

Internet Source

<1 %

47

awangjivi.com

Internet Source

<1 %

48

windows8manual.blogspot.co.id

Internet Source

<1 %

49

ebooktheory.blogspot.com

Internet Source

<1 %

50

olahdatajogjaa.blogspot.co.id

Internet Source

<1 %

51

digilib.unila.ac.id

Internet Source

<1 %

52

bse.sman5malang.sch.id

Internet Source

<1 %

53

aldysyafhidatiasuar.blogspot.com

Internet Source

<1 %

54

scundip.blogspot.com

Internet Source

<1 %

55

www.btu-online.de

Internet Source

<1 %

56

anzdoc.com

Internet Source

<1 %

57

staff.unila.ac.id

Internet Source

<1 %

58

tesis-ku.blogspot.com

Internet Source

<1 %

59

repository.unikom.ac.id

Internet Source

<1 %

60

perpustakaan.or.id

Internet Source

<1 %

Exclude quotes On

Exclude matches Off

Exclude bibliography On